

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Sciences du langage** Spécialité Français Langue
Étrangère

Arrêté ministériel : 7 août 2006

Présentée par

Thi Thu Hoai TRAN

Thèse dirigée par **Agnès TUTIN** et
codirigée par **Cristelle CAVALLA**

préparée au sein du **Laboratoire LIDILEM**
dans l'**École Doctorale Langues, littératures et sciences
humaines**

**Description de la phraséologie
transdisciplinaire des écrits scientifiques et
réflexions didactiques pour l'enseignement
à des étudiants non-natifs.
Application aux marqueurs discursifs**

Thèse soutenue publiquement le **11 décembre 2014**,
devant le jury composé de :

M. Alain KAMBER

Professeur, Université de Neuchâtel, Rapporteur

M. Jean-Marc MANGIANTE

Professeur, Université d'Artois, Rapporteur, Président

M. Serge VERLINDE

Professeur, Katholieke Universiteit Leuven, Examineur

Mme. Cristelle CAVALLA

Maître de conférences, Université Paris 3, Co-directrice

Mme. Agnès TUTIN

Professeur, Université Grenoble Alpes, Directrice

*Université Joseph Fourier / Université Pierre Mendès France /
Université Stendhal / Université de Savoie / Grenoble INP*



Remerciements

Cette thèse n'aurait jamais vu le jour sans les aides si précieuses de nombreuses personnes.

Mes premiers remerciements vont à Agnès Tutin et à Cristelle Cavalla d'avoir accepté d'encadrer cette thèse. Je leur adresse ma plus profonde reconnaissance, Sans elles, je n'aurais jamais mené à bien ce travail de recherche. Je les remercie de leur soutien sans faille, de leur rigoureuse intellectuelle, et de longues discussions vitaminées avec elles. Je les remercie d'avoir dirigé cette thèse avec beaucoup de responsabilité et de dévouement, de longues heures qu'elles m'ont consacré pour des re(re)lectures de cette thèse. Merci beaucoup à Agnès et à Cristelle de m'avoir formée, de m'avoir transmis votre passion pour la recherche et surtout merci à vous d'avoir toujours cru en moi tout au long de ces dernières années.

Je tiens à adresser tous mes remerciements les plus sincères aux membres du jury. Je voudrais remercier Monsieur Alain Kamber et Monsieur Jean-Marc Mangiante d'avoir accepté d'être rapporteurs dans mon jury. Je tiens à remercier Monsieur Serge Verlinde d'avoir accepté d'évaluer ce travail. Je souhaiterais leur exprimer ma reconnaissance pour de nombreuses suggestions qu'ils vont me donner pour faire avancer cette recherche.

J'exprime mes profonds remerciements à Marie-Paule Jacques pour son aide généreuse d'avoir passé tout le corpus sous Perl. Je la remercie de tous ses encouragements ces derniers temps. Mes remerciements chaleureux vont aussi à Achille Falaise pour m'avoir aidée à plusieurs reprises dans la constitution de notre corpus et dans la conception de Dicorpus.

Je souhaite remercier tous mes amis du laboratoire LIDILEM qui m'ont accueillie à bras ouverts et m'ont fait une vie de doctorante heureuse. Merci à Aurélie, Rozenn, Montiya, Tiphaine, Milie, Lulu, Mathieu, Sylvain, Isabelle, Isabelle, Julie, Manon, Eleni, Virginie, Laurence, etc. Je les remercie de leurs encouragements, de leurs messages d'amour, surtout aux moments les plus fastidieux en fin de thèse. Je les remercie des petites attentions qu'ils m'ont accordées, de petits câlins pour me rassurer. Je n'oublierai jamais toutes nos pauses en bas du D, nos midis au patio, nos soirées en ville et aussi des leçons sur « la faune et la flore », de petits cours de français, etc.

J'adresse mes remerciements les plus sincères à tous mes relecteurs dévoués et compétents qui ont accepté de relire ma thèse même au dernier moment. Merci à Rozenn, Véro, Milie, Lulu, Sylvain, Julie, Catherine, Amadou. Merci à Rozenn de ton amitié, de tes encouragements et de nombreuses relectures que tu m'as faites ces derniers temps. Merci à ma Véro de ton aide si précieuse, de tes relectures et des discussions vivifiantes avec toi. Merci à

Milie de tes petites attentions et de tes câlins rassurants. Merci à Lulu de tes bagels et de tous les petits messages. Merci beaucoup à Isabelle de ton aide si précieuse dans la constitution de cette bibliographie.

J'exprime ma reconnaissance à tous les membres du laboratoire Lidilem qui ont su toujours me motiver et m'encourager. Merci à Francis Grossmann, Laura Hartwell, Julie Sorba, Catherine Carras, Catherine Muller, Laura Abou Haida, je les remercie pour tout.

Mes remerciements vont à mes chers amis à la bibliothèque Sciences du Langage, j'ai passé deux années merveilleuses à vos côtés. Merci à Julie de m'avoir laissée travailler tranquillement ma thèse pendant ces derniers mois. Merci à Nadia de son partage. Merci à tous les moniteurs à la bibliothèque qui m'ont remplacée pour que je puisse travailler ma thèse.

Au niveau administratif, j'adresse mes remerciements sincères à Zohra et à Martine, elles sont pour moi les meilleures secrétaires du monde.

Ma profonde reconnaissance va à ma maman-poule-bác Régine et à mon papa-poule-bác Jean-Claude de tout ce qu'ils ont fait pour moi pendant toutes ces dernières années. Je les remercie de leur partage, de toutes les sorties culturelles, et de super bons repas et plein d'autres. Je les remercie de m'avoir transmis leurs connaissances et de petites attentions qu'ils m'ont accordées. Surtout je les remercie d'avoir pris soin de moi ces dernières années.

Je remercie la famille Préfasse qui m'a fait venir en France pour la première fois et m'a fait découvrir la France pendant toutes ces dernières années.

Tous mes remerciements vont à tous mes amis vietnamiens, de près ou de loin, qui s'inquiètent de l'avancement de mon travail de recherche. Merci à chị Kim, chị Thủy, chị Hà, chị Hiền, chị Linh, anh Dũng, chị Bái, chị Thúy, em Quỳnh, em Linh et beaucoup d'autres amis de votre soutien, de vos encouragements. Merci à Rui Yan de ton soutien moral si précieux.

Je voudrais remercier ma famille, mes parents de leurs encouragements, malgré la distance, qui m'ont été très significatifs. Je pense à ma petite maman qui a toujours su m'encourager même si elle a beaucoup souffert ces dernières années de mon absence.

Enfin, merci à toi Son, d'être toujours à mes côtés. Je te remercie de ton soutien, de tes encouragements et de ton amour.

Sommaire

LISTE DES TABLEAUX	IV
LISTE DES FIGURES	VI
INTRODUCTION	1
PARTIE 1 : OBJETS ET METHODOLOGIE DE RECHERCHE	7
CHAPITRE 1. ECRITS SCIENTIFIQUES ET PHRASEOLOGIE TRANSDISCIPLINAIRE SCIENTIFIQUE	8
1.1. Ecrit scientifique vs écrit universitaire : définition.....	8
1.2. Phraséologie transdisciplinaire scientifique : définition	12
1.3. Phraséologie transdisciplinaire scientifique : typologie.....	15
1.4. Intérêt pour le lexique transdisciplinaire scientifique	23
Conclusion.....	26
CHAPITRE 2. MARQUEURS DISCURSIFS	28
2.1. Intérêt pour les marqueurs discursifs	29
2.2. Marqueurs discursifs : enseignement/apprentissage en FLE	33
2.3. Marqueurs discursifs : typologie.....	46
Conclusion.....	51
CHAPITRE 3. METHODOLOGIE	53
3.1. Une approche inscrite dans la linguistique de corpus.....	53
3.2. Constitution du corpus	54
3.3. Repérage des séquences polylexicales	61
3.4. Approches adoptées.....	74
Conclusion.....	76
PARTIE 2 : ANALYSES LINGUISTIQUES	77

CHAPITRE 4. TYPOLOGIE DES SEQUENCES LEXICALISEES A FONCTION DISCURSIVE	78
4.1. Principes sous-jacents de notre typologie	79
4.2. Typologie des MD.....	81
4.3. Marqueurs de reformulation	93
Conclusion.....	107
 CHAPITRE 5. MARQUEURS DE REFORMULATION : ANALYSE SYNTACTICO-SEMANTIQUE	 108
5.1. Principes de la grille d'analyse	108
5.2. Paramètres de la grille d'analyse	110
5.3. Clarifier	120
5.4. Récapituler	155
5.5. Reconsidérer.....	172
Conclusion.....	200
 PARTIE 3 : RÉFLEXIONS DIDACTIQUES	 203
 CHAPITRE 6. PROPOSITIONS DIDACTIQUES POUR L'ENSEIGNEMENT/ APPRENTISSAGE DE SEQUENCES LEXICALISEES A FONCTION DISCURSIVE.....	 204
6.1. Corpus de mémoires d'apprenants	205
6.2. Linguistique de corpus et enseignement/apprentissage du français.....	212
6.3. Réflexions didactiques : vers des propositions.....	219
Conclusion.....	276
 CONCLUSION GENERALE.....	 278
 BIBLIOGRAPHIE	 283
 INDEX.....	 293
 ANNEXE	 295
Annexe 1: Typologie des séquences lexicalisées à fonction discursive	296

Annexe 2 : Liste des revues en SHS	305
Annexe 3 : Corpus de mémoires d'apprenants vietnamophones	324

Liste des tableaux

Tableau 2.1 : Propriétés principales des MD.....	33
Tableau 2.2 : Erreurs fréquentes de l'utilisation des MD chez les locuteurs non natifs	37
Tableau 2.3 : Liste des méthodes de FLE consultées.....	39
Tableau 2.4 : Livres de grammaire de FLE consultés	42
Tableau 3.1 : Liste des disciplines en SHS dans notre corpus	55
Tableau 3.2 : Sélection des séquences polylexicales : critère de fréquence et de répartition.....	72
Tableau 3.3 : Sélection des séquences polylexicales : critères de fréquence, de répartition et de pertinence	74
Tableau 4.1 : Différentes combinaisons de différentes propriétés (Roulet, 1987 : 120).....	102
Tableau 4.2 : Des opérations de reformulation (Rossari, 1994 : 22).....	104
Tableau 4.3 : Liste des marqueurs de reformulation dans notre corpus SHS.....	106
Tableau 5.1 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>c'est-à-dire</i>	131
Tableau 5.2 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>à savoir</i>	138
Tableau 5.3 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>autrement dit</i>	147
Tableau 5.4 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>en d'autres termes</i>	153
Tableau 5.5 : Comparaison des caractéristiques de <i>c'est-à-dire</i> , <i>à savoir</i> , <i>autrement dit</i> et <i>en d'autres termes</i>	155
Tableau 5.6 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>en somme</i>	161
Tableau 5.7 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>en résumé</i>	167
Tableau 5.8 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>en un mot</i>	171
Tableau 5.9 : Comparaison des caractéristiques de <i>en somme</i> , <i>en résumé</i> et <i>en un mot</i>	172
Tableau 5.10 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>au total</i>	179
Tableau 5.11 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>somme toute</i>	183
Tableau 5.12 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>en définitive</i>	189
Tableau 5.13 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>en fin de compte</i>	194
Tableau 5.14 : Récapitulatif des propriétés syntaxiques et sémantiques de <i>au fond</i>	198
Tableau 5.15 : Comparaison des caractéristiques de <i>au total</i> , <i>somme toute</i> , <i>en définitive</i> , <i>en fin de compte</i> et <i>au fond</i>	200

Tableau 6.1 : Corpus de mémoires d'apprenants vietnamophones.....	206
Tableau 6.2 : Fréquence des marqueurs de reformulation dans le corpus d'apprenants vietnamophones.....	210

Liste des figures

Figure 1-1 : Extrait du schéma collocationnel pour la notion CONSEQUENCE (D'après Pecman, 2004 : 394)	19
Figure 1-2 : Accès onomasiologique dans LEAD (d'après Granger et Paquot, 2010 : 323).....	23
Figure 2-1 : Place des cadratifs parmi l'ensemble des constituants (Charolles, 2003)	48
Figure 3-1 : Structure d'un document TEI Lite.....	60
Figure 3-2 : Distributionnal categorie (d'après Granger & Paquot, 2008 : 39).....	62
Figure 3-3 : Liste des disciplines dans le corpus SHS – Interface Scientext (corpus Intranet).....	65
Figure 3-4 : Sélectionner des articles en linguistique.....	65
Figure 3-5 : Liste des articles sélectionnés en linguistique	66
Figure 3-6 : Différents modes de recherche dans Scientext	66
Figure 3-7 : Visualisation en concordancier des occurrences des locutions adverbiales en Linguistique.....	67
Figure 3-8 : Visualisation en lemme des résultats des locutions adverbiales en linguistique	68
Figure 3-9 : Répartition en partie textuelle des résultats des locutions adverbiales	68
Figure 3-10 : Confrontation des résultats dans Access pour des locutions adverbiales	69
Figure 4-1 : Exemple de l'utilisation des séquences polylexicales à fonction discursive dans un article scientifique	81
Figure 4-2 : Comparaison de l'utilisation des marqueurs de reformulation dans le corpus SHS et dans le corpus Emolex.....	94
Figure 6-1 : Concordance de <i>autrement dit</i> dans le corpus d'apprenants	207
Figure 6-2 : Navigation entre les occurrences et le contexte large.....	208
Figure 6-3 : Base de données des séquences lexicalisées à fonction discursive dans Dicorpus – Entrée onomasiologique	232
Figure 6-4 : Dicorpus - Entrée sémasiologique.....	232
Figure 6-5 : Dicorpus - Exemples validés sur le corpus Scientext.....	233
Figure 6-6 : Dicorpus - Contexte large de <i>en somme</i>	233
Figure 6-7 : Schéma des stratégies de lecture <i>top-down</i> et <i>bottom-up</i>	240
Figure 6-8 : Schéma d'analyse - Fiche pédagogique 1	257
Figure 6-9 : Schéma d'analyse - Fiche pédagogique 2	265

Figure 6-10 : Schéma d'analyse - Fiche pédagogique 3	270
Figure 6-11 : Schéma d'analyse - Fiche pédagogique 4	275

Introduction

Notre travail de recherche porte sur la phraséologie transdisciplinaire scientifique, et plus précisément, sur les séquences lexicalisées à fonction discursive, afin de proposer une méthodologie d'aide à la rédaction scientifique aux étudiants non natifs du français. Les programmes d'échanges européens et internationaux étant de plus en plus nombreux, nous assistons à une forte croissance de la « mobilité universitaire ». Les effectifs d'étudiants non francophones venus en France pour leurs études ont fortement augmenté au cours des dernières années. Ces étudiants sont issus de systèmes académiques très diversifiés et disposent d'un bagage linguistique différent. Ils doivent tous se confronter à de nombreuses difficultés d'intégration sociale, institutionnelle et scientifique, et se familiariser avec les normes d'une nouvelle culture disciplinaire, suivre des contenus universitaires en langue étrangère, s'approprier de nouvelles pratiques d'écritures, etc. Face aux besoins scientifiques et linguistiques de ces étudiants, plusieurs établissements déploient des moyens comme des formations à la méthodologie universitaire afin de concevoir des dispositifs d'enseignement à l'intention des étudiants étrangers et faciliter ainsi leur « intégration universitaire ».

Notre recherche s'inscrit dans une démarche de *Français sur Objectifs Universitaires* (FOU) (Cf. Mangiante & Parpette, 2011, plus précisément dans sa dimension linguistique qui est parfois jugée comme le pavé manquant du FOU au profit de la dimension méthodologique. Face à une nouvelle culture disciplinaire, les apprenants doivent assimiler les règles linguistiques et textuelles du genre des écrits universitaires. Nous nous intéressons particulièrement aux étudiants locuteurs non natifs du français de 2^e ou 3^e cycles dans les universités françaises ou francophones qui ont à rédiger un écrit scientifique (rapport de recherche, mémoire, thèse ou article de recherche). Il s'agit d'un public de niveau avancé dans cette langue, mais qui éprouve encore quelques difficultés rédactionnelles quant à la structure du discours, ou l'utilisation en contexte des combinaisons de mots. Les difficultés des étudiants se situent essentiellement à l'intersection entre la langue courante et les langues

spécialisées, c'est-à-dire que les moyens d'expression appartiennent à la langue générale, mais sont utilisés à des fins scientifiques, ce qui explique l'intérêt que nous portons à la fois à la phraséologie transdisciplinaire scientifique, mais aussi aux marqueurs discursifs. Par ailleurs, par ce travail, nous souhaitons développer les compétences discursives des étudiants en commençant par mettre à leur disposition des moyens lexicaux qui les aident à mieux organiser leur texte et à produire un discours cohérent.

Notre travail se situera au croisement de plusieurs domaines : la sémantique lexicale, la didactique du lexique spécialisé et la linguistique de corpus associée au traitement automatique du langage (TAL). Ce travail exploite les techniques du TAL au service de la didactique, en recourant aux méthodes de la linguistique de corpus, encore peu exploitée dans le champ didactique francophone, pour l'enseignement/apprentissage du lexique transdisciplinaire des écrits scientifiques.

1. Intérêt de l'utilisation du corpus dans le domaine du Français Langue Etrangère (FLE)

Notre approche linguistique s'inscrit dans le courant de la linguistique de corpus, dont les origines, au moins dans la tradition anglo-saxonne, sont souvent liées à des préoccupations de didactique des langues. Le *corpus BROWN* est l'un des premiers corpus électroniques à avoir vu le jour avec environ un million de mots, élaboré par Kučera et Francis en 1967, ainsi que *The Bank of English* (lancé par Cobuild et l'Université de Birmingham en 1991). On mentionnera également, dans les années 90, le très gros corpus *BNC British National Corpus* avec 100 millions de mots étiquetés, venant de différentes sources écrites et orales, a été créé dans le cadre de l'initiative d'un consortium académico-industriel composé des membres de Oxford University Press, Addison-Wesley Longman. En France, l'idée d'un corpus représentatif de la langue reste peu développée malgré de gros projets assez novateurs et l'initiative Corpus de référence lancée récemment par l'Institut de Linguistique Française¹. Dans les années 50, le *Français Fondamental* de Gougenheim *et al.* (1964) a été lancé dans le but de constituer un lexique de mots de base à partir d'enquêtes orales, puis dans les années 60, *Frantext* a fait son apparition avec plus de 100 millions de mots et reste le plus gros corpus de littérature composé essentiellement des romans depuis le XVI^e siècle. Malgré l'intérêt des corpus annotés comportant des éléments phraséologiques, ce type de ressources n'a pas été énormément développé pour le traitement des séquences lexicalisées à fonction

¹ <http://www.ilf.cnrs.fr/spip.php?rubrique95>

discursive et les applications didactiques². C'est pourquoi nous avons souhaité nous intéresser de manière approfondie à cette perspective qui jusqu'alors a été peu explorée.

L'introduction du corpus en classe de langue, très en vogue dans le monde anglo-saxon, a commencé dès les années 90 avec l'approche *data driven-learning* de Johns (1991). Cette approche est plus centrée sur l'apprenant et correspond bien à l'approche constructiviste de Vygotsky (1896-1934). Elle privilégie l'autonomie d'apprentissage des étudiants tout en les laissant relever eux-mêmes des règles à partir d'un corpus ou d'un concordancier. Pourtant, en France, les enseignants de FLE manifestent encore des réticences face à l'introduction d'un tel outil dans l'apprentissage des langues même si de nombreux travaux prêtent une attention croissante à l'utilisation des corpus, et en particulier à l'usage des concordanciers dans les classes de langue pour la prise de conscience des phénomènes linguistiques (Boulton & Tyne, 2014 ; Granger & Paquot, 2009). Dans ce travail, nous suggérons des pistes d'utilisation pour les enseignants dans l'apprentissage sur corpus.

Nous nous efforçons de relier les descriptions linguistiques faites sur le corpus pour alimenter par la suite les activités didactiques. En particulier, les analyses linguistiques et les activités didactiques sont effectuées sur un corpus annoté morphologiquement et syntaxiquement grâce aux outils Traitement Automatique des Langues (TAL). L'utilisation d'un corpus d'écrits scientifiques est très pertinente pour un public universitaire qui a un réel besoin de s'immerger dans un « bain linguistique » pour se familiariser avec ce lexique spécifique ainsi que les règles de ce genre d'écrits. Nous souhaitons développer la prise de conscience des phénomènes linguistiques chez les étudiants grâce au contact avec les écrits scientifiques.

2. Intérêt du travail pour les marqueurs discursifs

Ce travail se situe dans une conception étendue du domaine de la phraséologie comme évoqué par Legallois & Tutin (2013). Nous nous intéressons aux éléments phraséologiques spécifiques, plus précisément les unités polylexicales qui jouent un rôle d'organisation du texte et de structuration du discours, que nous appelons séquences lexicalisées à fonction discursive. Ces unités lexicales ont fait l'objet de nombreuses études dans la littérature et portent différentes appellations dont *particules énonciatives* (Fernandez, 1994), *connecteurs* (Riegel *et al.*, 2009), *marqueurs de discours* (Péry-Woodley, 2005 ; Ho-Dac *et al.*, 2012),

² Leprojet ANR Annodis, qui traite de ces questions, a plutôt une orientation linguistique, Ressources Annodis *Annotation Discursive Corpus de français écrit enrichi d'annotations discursives* : <http://redac.univ-tlse2.fr/corpus/annodis/>

marqueurs de glose (Steuckardt & Niklas-Salminen, 2005), *marqueurs discursifs* (Dostie & Pusch, 2007 ; Paillard & Vu, 2012), etc. Nous retiendrons dans notre travail les termes *marqueurs discursifs* (désormais MD), uniquement pour des raisons de brièveté. Nous nous intéressons au fonctionnement de ces éléments phraséologiques dans un genre spécifique, des écrits scientifiques. Nous les intégrons dans la phraséologie transdisciplinaire scientifique, phraséologie qui n'est pas absente de la langue générale, mais qui prend des acceptations spécifiques dans les écrits scientifiques et y est particulièrement surreprésentée (Drouin, 2007 ; Tutin, 2007 b ; Paquot, 2010).

Ces MD posent des difficultés aussi bien aux enseignants de FLE qu'aux apprenants non natifs du français en raison de leurs complexités d'ordre morphosyntaxique, syntaxique et sémantique. En outre, ces expressions polylexicales ne font pas encore l'objet d'un enseignement systématique dans les manuels de FLE. Par ailleurs, à notre connaissance, il existe encore peu de travaux dans la littérature qui cherchent à combiner ces descriptions linguistiques avec des applications didactiques. Ce travail propose un premier essai d'introduction du corpus électronique dans la classe de langue pour enseigner les phénomènes linguistiques qui nous intéressent.

Nos objectifs de recherche sont à la fois d'ordre linguistique et didactique. Nous souhaitons d'abord proposer une modélisation syntaxique et sémantique de la phraséologie scientifique et mettre au point des dispositifs d'aide à l'écriture basés sur l'accès lexical. Nous nous intéresserons à un sous-ensemble de MD, les marqueurs de reformulation, et procéderons à une analyse détaillée des marqueurs les plus fréquents dans notre corpus. Cette analyse nous permettra de relever les régularités syntaxiques et sémantiques des MD. En ce qui concerne les objectifs didactiques, nous voudrions mettre en place une méthodologie pour l'enseignement/apprentissage de ces unités lexicales, et évaluer la pertinence de la linguistique de corpus pour les activités de didactique du lexique spécialisé. Pour ce faire, nous allons tirer profit des analyses linguistiques réalisées. En effet, la typologie des séquences lexicalisées à fonction discursive nous servira par la suite dans la conception d'activités pour les activités didactiques.

Nous formulons ici l'hypothèse qu'il existe un lien étroit entre le fonctionnement syntaxique et sémantique des MD. Autrement dit, les valeurs sémantiques des MD ne peuvent être définies seulement par leur sens, mais aussi par la relation entre les deux séquences établie par le MD. Les spécificités syntaxiques et sémantiques de chaque MD permettront de distinguer un MD d'un autre souvent considéré comme synonyme.

La problématique qui sous-tend cette recherche est relative à un modèle d'analyse des MD et à la transposition didactique des éléments linguistiques. Nous pouvons l'énoncer sous la forme des questions suivantes : quels sont les paramètres à prendre en compte pour une analyse des MD ? Quels sont les rôles des séquences lexicalisées à fonction discursives dans la structuration du discours ? Existe-t-il une approche idéale pour l'enseignement/apprentissage de ces unités lexicales ? Comment aider les apprenants à réutiliser ces moyens lexicaux dans leurs écrits ?

3. Structure de cette étude

Ce travail est composé de trois parties. La première partie décrit les objets de recherche et la méthodologie. La deuxième partie porte sur l'analyse linguistique des séquences lexicalisées à fonction discursive, et des marqueurs de reformulation en particulier. La troisième partie est consacrée au développement des réflexions didactiques autour de l'enseignement/apprentissage de ces unités lexicales à l'appui d'un corpus.

Le chapitre 1 constitue une entrée en matière où nous faisons l'état de l'art de la phraséologie transdisciplinaire dans les écrits scientifiques. Nous définissons avant tout ce genre spécifique, celui des écrits scientifiques et le distinguons des écrits universitaires. Ces écrits rédigés par des experts dans le domaine en question sont considérés comme un modèle pour les étudiants. Nous expliquerons par la suite l'importance d'un enseignement/apprentissage du lexique transdisciplinaire dans les écrits scientifiques pour développer les compétences linguistiques et discursives des apprenants.

Le chapitre 2 explicite notre objet de recherche, les séquences lexicalisées à fonction discursive ou MD polylexicaux. Nous relevons dans ce chapitre les différentes propriétés des MD qui posent souvent des difficultés aux étudiants non natifs du français. Nous présentons ensuite quelques typologies des MD qui nous inspirent par la suite dans ce travail.

Le chapitre 3 présente la méthodologie adoptée. Nous nous appuyons sur une étude de corpus et sur une analyse empirique des marqueurs de reformulation. Nous décrivons le rôle de l'étude de corpus pour cette recherche et les critères de sélection des séquences polylexicales.

Le chapitre 4 met en place une typologie des séquences lexicalisées à fonction discursive, et aborde les marqueurs (exemplification, reformulation, etc.) qui sont surreprésentés aux écrits scientifiques en SHS.

Le chapitre 5 est consacré à la modélisation syntaxique et sémantique des 12 marqueurs de reformulation que nous regroupons en trois sous-groupes, les marqueurs servant à clarifier, à récapituler, et à reconsidérer. Ces marqueurs sont analysés du point de vue syntaxique (portée, position) et sémantique, ce qui permet de relever les régularités syntaxiques et sémantiques des MD.

Le chapitre 6 fait écho au chapitre 5 à travers la proposition d'activités didactiques en lien avec les spécificités linguistiques des éléments décrits. Nous proposons une entrée méthodologique pour l'enseignement/apprentissage des marqueurs discursifs polylexicaux, mais également pour les autres unités lexicales.

PARTIE 1 :
OBJETS ET METHODOLOGIE DE
RECHERCHE

Chapitre 1.

Ecrits scientifiques et phraséologie transdisciplinaire scientifique

Dans ce chapitre, nous allons circonscrire notre domaine de travail. Nous allons décrire tout d'abord le genre de texte qui nous intéresse, en particulier les écrits scientifiques. Ensuite, nous allons définir le lexique dans lequel s'intègre notre objet de recherche, le lexique transdisciplinaire des écrits scientifiques. Nous finirons par un récapitulatif des erreurs d'apprenants dans la rédaction de leurs productions écrites pour justifier le choix de notre sujet.

1.1. Ecrit scientifique vs écrit universitaire : définition

Dans notre travail, nous allons partir des analyses syntactico-sémantiques des éléments phraséologiques des écrits scientifiques pour ensuite les appliquer à l'écriture universitaire scientifique, la perspective étant de développer une méthodologie d'aide à la rédaction scientifique pour les étudiants non natifs du français. Il nous paraît donc indispensable de définir le lien entre écrits scientifiques et écrits universitaires. Dans quelle mesure un de ces deux genres de textes (scientifique ou universitaire) inclut l'autre dans son champ ? Quelles en sont les principales différences ? Pour répondre à ces questions, nous cherchons à délimiter la frontière entre ces deux genres.

1.1.1. Écrits scientifiques

1.1.1.1. Définition et propriétés

Il nous semble intéressant, de prime abord, d'expliciter la notion de *scientifique* souvent liée par abus aux sciences dures, exactes ou appliquées (comme les mathématiques, la biologie, la chimie, la physique, l'informatique). En réalité, de nombreuses recherches en

linguistique et en FOS Français sur Objectifs Spécifiques (Phal, 1971 ; Eurin Balmet & Henao de Legge, 1992 ; Pecman, 2004) traitent du français scientifique comme le français appartenant aux disciplines des mathématiques, de la chimie et de la physique. Dans notre étude, nous avons une définition plus large de la notion de *scientifique* : nous considérons ce qui relève du scientifique comme l'« ensemble structuré de connaissances [...] obéissant à des lois objectives (ou considérées comme telles) et dont la mise au point exige systématisation et méthode » (d'après le TLFi³), c'est-à-dire comme un cadre théorique de recherche, des hypothèses, la problématique, des réponses à la problématique. Cette notion englobe donc tant les sciences dures et exactes que les sciences humaines et sociales (désormais SHS).

D'après Berthelot (2003), un texte scientifique comporte trois critères : « une intention de connaissance explicite de l'auteur, un apport de connaissance reconnu par une communauté savante, l'inscription dans un espace de publication identifiable comme scientifique » (*ibid.* : 33). Ces critères s'articulent dans le double mouvement de la synchronie et de la diachronie. D'une part, un texte scientifique répond aux normes de présentation et d'exposition du moment de sa publication tout en assumant la validité de son apport de connaissances. D'autre part, quand il s'inscrit dans le travail de réflexivité mené à long terme dans les disciplines, il doit résister aux procédures d'évaluation et de révision d'une communauté savante selon les normes croisées de recevabilité et de scientificité. Au-delà des variations culturelles et disciplinaires, l'écrit scientifique relève d'une structure commune qui se fonde sur le système de la preuve et la prétention à la rationalité. L'écrit scientifique s'oppose au texte littéraire en ce qu'il utilise un langage formulaire contrairement à un langage poétique, métaphysique, littéraire d'où l'importance d'un enseignement des formules, des phrasèmes spécifiques et fortement représentés dans ce type d'écrits.

1.1.1.2. Discours scientifique

Analyser les écrits scientifiques nous amène à analyser le discours scientifique, car celui-ci « procède de l'écrit et des possibilités offertes par ce dernier afin de fixer la pensée et de la mettre en débat » (Berthelot, 2003 : 21). Le discours scientifique est le « discours produit dans le cadre de l'activité de recherche à des fins de construction et de diffusion du savoir » (Rinck, 2010 : 428), et « le seul authentique reflet de l'activité scientifique » (Jacobi, 1999 : 26). La notion de discours scientifique varie selon les auteurs, ce qui offre différentes classifications de cette notion. D'après Loffler-Laurian (1983), le terme « discours

³ Site du TLFi Trésor de la Langue Française Informatisée : <http://atilf.atilf.fr/>

scientifique » est « une appellation commode pour désigner une variété de discours ». Il renvoie donc à « l'ensemble des textes écrits et des productions orales ayant un contenu dit scientifique, c'est-à-dire lié à la recherche, l'enrichissement et la diffusion des connaissances sur la nature et le fonctionnement du monde minéral, végétal, humain, etc. » (*ibid.* : 9).

Le discours scientifique se caractérise par l'apparition obligatoire des terminologies dans les textes scientifiques que Jacobi (1999) entend par sa propriété « ésotérique ». En outre, par son analyse minutieuse du français technoscientifique, Kocourek (1991) a précisé que les écrits scientifiques comportent un sous-langage spécifique, caractérisé par des traits textuels, pragmatiques, syntaxiques et lexicaux remarquables. Cette caractéristique pose sans doute des difficultés pour les non-spécialistes du domaine en question.

Plusieurs typologies du discours scientifique ont vu le jour. Selon Poudat (2006), le discours scientifique s'inscrit dans la perspective de communication scientifique et se divise en deux sous-pratiques selon la nature du public visé : la vulgarisation scientifique et le discours académique. Tandis que la vulgarisation scientifique cherche à communiquer la recherche au grand public et se situe plutôt au niveau du discours journalistique, le discours scientifique académique présente « la particularité de faire coïncider le public de ses producteurs et celui de ses consommateurs » (*ibid.* : 51). Dans notre travail, nous nous intéressons uniquement à cette deuxième catégorie, c'est-à-dire aux discours qui sont écrits dans des contextes précis et ne sont diffusés que dans des cercles précis et déterminés d'utilisateurs. Il s'agit en fait des discours scientifiques qui sont publiés dans des revues savantes dans l'objectif de produire de la connaissance. Ils correspondent aux *discours scientifiques primaires* dans la typologie proposée par Jacobi (1999) ou au *discours scientifique spécialisé* dans la classification de Löffler-Laurian (1983). Le discours scientifique se réalise à travers différents genres qui soumettent la forme et le contenu aux variations culturelles et disciplinaires : comptes rendus, articles, conférences, posters, etc..

Nous allons travailler dans cette recherche sur les articles de recherche en SHS qui ont été rédigés par des chercheurs chevronnés du domaine dans l'intention de communiquer leur connaissance et publiés dans des revues spécialisées. Ces écrits comportent donc des spécificités du discours scientifique (structuration, langage formulaire, énonciation, système de raisonnement discursif). Nous allons décrire en détail le genre de l'article de recherche ultérieurement.

1.1.2. Écrits universitaires

Il existe différents types d'écrits universitaires comme le rapport, le compte rendu de lecture, le dossier, le mémoire, etc. L'écrit universitaire reste le moyen d'évaluation privilégié des études supérieures dans l'intention de terminer une formation ou d'obtenir un diplôme. Il constitue l'aboutissement de plusieurs mois ou plusieurs années de travail. L'apprenant doit en maîtriser les normes tant sur la forme de l'écrit universitaire, comme les critères de forme (orthographe, organisation de l'espace, typographie, etc.) ou la clarté des constructions phrastiques, que sur le fond, à savoir la cohérence sémantique, la pertinence des méthodes de travail, la logique et la prise en compte d'une bonne structure du texte. Il est demandé à l'étudiant à travers sa production de manifester sa capacité à assimiler des connaissances ainsi que d'acquérir d'autres compétences scripturales, par exemple s'exprimer de façon claire et rigoureuse, communiquer et exprimer son point de vue et son raisonnement, etc. Il doit également savoir poser une problématique, des hypothèses, analyser des informations et des documents, organiser et argumenter ses idées, synthétiser, tout en ouvrant sur de nouvelles questions.

Comme nous l'avons précisé préalablement, nous nous intéressons tout particulièrement aux disciplines de SHS pour plusieurs raisons. La première raison est liée à une demande importante de la part des étudiants venant de ces disciplines, auxquels il est davantage demandé d'argumenter que dans les disciplines de sciences exactes. En effet, ces disciplines en SHS sont caractérisées par la présence d'un métadiscours complexe, et sont plus souvent susceptibles d'interprétations divergentes (Biber, 2006) que dans le domaine des sciences purement formelles comme les mathématiques, ou le domaine expérimental comme les sciences physiques, les sciences naturelles, les sciences biologiques, etc. Ces dernières sont caractérisées par des unités observables, identifiables, mesurables. La rédaction d'un mémoire ou d'une thèse dans ces domaines est, nous semble-t-il, plus aisée que dans le domaine des lettres, arts, sciences humaines, etc. Dans ces domaines, c'est la langue qui véhicule en quelque sorte la science. C'est pourquoi les besoins linguistiques sont les plus manifestes (Bard, 2007), car il s'agit de « besoins spécifiques au niveau de la structuration et de la construction des écrits, autant au niveau du contenu que du contenant » (*ibid.* : 41).

La deuxième raison est relative aux spécificités des écrits scientifiques. Rédiger un écrit universitaire constitue une tâche difficile pour les étudiants, particulièrement les étudiants non natifs qui rédigent en langue étrangère. Plusieurs recherches menées au LIDILEM Laboratoire Linguistique et Didactiques des Langues Etrangères et Maternelles

mettent en évidence l'importance de l'enseignement du lexique et de la rédaction d'écrits universitaires (Cavalla, 2007, 2009 ; Carras & Cavalla, 2007). Outre les normes et les dispositions réglementaires à respecter, l'étudiant doit manifester des aptitudes intellectuelles, une maîtrise des méthodes employées, de la rigueur scientifique, etc. Les difficultés rencontrées par les étudiants sont aussi d'ordre culturel, car les exigences universitaires diffèrent d'un pays à l'autre. Les écrits universitaires français ont donc leurs propres spécificités (Cislaru, Claudel & Vlad, 2009 ; Eckenschwiller, 1994 ; N'Da, 2007). Comme le mentionne Cavalla (2007), les étudiants doivent manipuler simultanément deux savoirs : le français en tant que langue étrangère et le savoir scientifique dont ils ne sont pas encore experts. Leurs besoins se situent sur quatre dimensions, à savoir les dimensions scientifique, méthodologique, terminologique et linguistique. À la dimension scientifique correspond le savoir disciplinaire, la dimension méthodologique aborde la structure des écrits universitaires, la dimension terminologique traite du lexique savant ; enfin, la dimension linguistique étudie les structures qui contribuent au sens.

Pour répondre aux questions que nous avons posées antérieurement, les écrits universitaires avec leurs caractéristiques constituent un sous-domaine des écrits scientifiques, genre large et englobant. C'est pourquoi, dans notre travail, nous allons recourir aux écrits scientifiques, essentiellement des articles de recherche, qui nous serviront en quelque sorte de **modèle** vers lequel les apprenants peuvent tendre pour rédiger leur écrit universitaire.

1.2. Phraséologie transdisciplinaire scientifique : définition

La phraséologie est particulièrement présente dans les écrits scientifiques et a fait l'objet de nombreux travaux de recherche en anglais (Howarth, 1998 ; Gledhill, 2000 ; Granger & Meunier, 2008) comme en français (Tutin, 2007 a, 2008 ; Pecman, 2004). Plusieurs auteurs s'accordent pour affirmer l'existence d'un vocabulaire non technique, non soumis aux variations disciplinaires, mais appartenant à plusieurs domaines et qu'il est indispensable de maîtriser dans la rédaction scientifique et universitaire (Drouin, 2007 ; Tutin, 2007 b ; Paquot, 2010).

Phal (1971) a été le pionnier dans l'analyse de la langue générale scientifique en français en élaborant à partir d'un corpus de manuels scientifiques le *Vocabulaire Général d'Orientation Scientifique* (VGOS). Il l'a défini comme un lexique « commun à toutes les

spécificités » (*ibid.* : 9). Se situant dans la même lignée que *Français Fondamental*⁴ de Georges Gougenheim *et al.* (1964), le VGOS de Phal a pour but de faciliter l'accès aux études scientifiques en langue française auprès des étudiants étrangers (*ibid.* : 7). Il a établi une liste de mots qui sera au service des professeurs de langue de spécialité afin d'aider les étudiants à maîtriser la langue scientifique générale. L'extraction du VGOS s'est faite à partir d'un corpus de 1 794 500 mots extraits de manuels scientifiques du second cycle de disciplines en sciences exactes, dont les mathématiques, les sciences physiques (physique et chimie) et les sciences naturelles (sciences de la vie et de la terre). Il se présente comme une liste alphabétique suivie d'indices de fréquence et d'autres éléments syntaxiques et sémantiques.

Se situant dans la perspective des didactiques des langues, Chetouani (1997) a mis en place un « vocabulaire général de communication pédagogique dans le domaine scientifique en français langue étrangère » (*ibid.* : 11), nommé *Vocabulaire général d'enseignement scientifique* VGES. Le VGES résulte d'un travail minutieux effectué à partir d'un corpus oral et d'un corpus écrit, dont une grande partie est consacrée au corpus oral de 24 h de cours enregistrés dans des classes du secondaire, essentiellement en sciences dures (physique, chimie), les sciences sociales (droit, économie) ne représentant qu'une petite partie du corpus (4 h d'enregistrement au total). Pour compléter la description du corpus oral de physique, Chetouani l'a mis en confrontation avec d'autres corpus ayant un mode de formulation de nature différente. Le VGES recense 811 termes pleins classés par ordre alphabétique et est accompagné d'indications sur la classe grammaticale, la fréquence, le rang et l'appartenance (au FF ou au VGOS). Il nous paraît dommage qu'une étude linguistique plus poussée ne soit pas développée ni pour le VGOS ni pour le VGES. En effet, le traitement sémantique n'est pas développé, ce qui ne permet pas une approche par concept. Nous allons revenir sur cette question dans le chapitre 6 consacré aux réflexions didactiques.

Ayant également pour objectif d'apporter aux étudiants étrangers un bagage linguistique nécessaire à leurs études à l'université, les travaux de Coxhead (2000) se situent dans la perspective de l'enseignement du vocabulaire en anglais pour le programme *English for Academic Purposes* (ou l'anglais pour objectifs universitaires). Elle met en place une liste de 2 000 mots de la langue universitaire ou *Academic Word List (AWL)*. Cette liste est extraite d'un corpus de 3 500 000 mots constitué de revues savantes, d'articles tirés du Web, de

⁴ *Le Français Fondamental* FF paru en 1954, refondu et augmenté en 1964. La recherche vise à isoler un lexique de base à partir de la méthode quantitative. La liste publiée comme FF recense de 1 475 mots à l'intention des concepteurs de méthodes pour débutants en Français Langue Etrangère et extraite d'un corpus de 31 200 items, soit 8 000 mots différents composés de 163 conversations sur des sujets très variés.

documents sélectionnés dans *Brown Corpus*⁵, *LOB corpus*⁶ et de manuels de quatre thématiques, dans lesquels 28 disciplines sont représentées au total, y compris les arts, le commerce, le droit et les sciences. *The Academic Word List* résulte d'une étude statistique complexe afin de sélectionner les termes les plus fréquents et représentatifs du discours scientifique. Cette liste est organisée par niveaux de fréquence et groupée en 570 familles de mots réparties en dix sous-listes par ordre de fréquence. Par exemple, la première sous-liste se compose des familles de mots les plus fréquentes du corpus comme *analyse, approach, concept, context, formula, method etc.* Coxhead ne procède pas à l'analyse des unités polylexicales. À l'instar des travaux du *Français Fondamental* de Gougenheim *et al.* (1964), Phal, Coxhead et Chetouani se basent tous sur des critères lexicométriques, des indices de fréquence pour proposer des listes de mots pleins et des mots grammaticaux, mais ils n'effectuent pas encore un classement notionnel. À la différence de ces auteurs, Pecman (2004) et Tutin (2007 a, 2007 b) se basent sur des critères de combinatoire lexicale pour établir un classement notionnel des unités phraséologiques, ce qui correspond à notre visée didactique.

Toujours pour des visées didactiques et se situant dans la même lignée que Phal (*op. cit.*), Pecman (2004) s'intéresse aux unités phraséologiques de la *Langue Scientifique Générale* dans une perspective contrastive anglais-français qu'elle décrit à partir d'un corpus en physique, en chimie et en biologie. Elle définit la *Langue Scientifique Générale* comme une :

pratique langagière spécifique à une communauté de discours composée de chercheurs en sciences exactes dont les objectifs communicatifs poursuivis émanent des préoccupations partagées par des scientifiques à travers le monde et indépendamment de leurs spécificités disciplinaires (*ibid.* : 128)

Même si la recherche de Pecman porte uniquement sur les sciences exactes, l'approche notionnelle proposée paraît intéressante. Nous allons y revenir dans la partie suivante portant sur différentes typologies de la phraséologie.

Dans la perspective de développer des outils d'aide à la rédaction, Tutin (2007 b) a proposé une ébauche de traitement sémantique du *lexique transdisciplinaire des écrits scientifiques*, qu'elle décrit comme :

⁵ *Brown Corpus*, élaboré en 1967 par Kučera et Francis de Brown University, est l'un des premiers corpus électroniques de taille importante à avoir vu le jour. Il recueille des textes sélectionnés par échantillonnage parmi quinze genres textuels différents (articles de presse, textes religieux, textes gouvernementaux, science-fiction, etc.)

⁶ *LOB corpus* ou *Lancaster-Oslo-Bergen corpus*, élaboré dans les années 1970 à l'issue de la collaboration de University of Lancaster, University of Oslo et Norwegian Computing Center for the Humanities.

le lexique partagé par la communauté scientifique mis en œuvre dans la description et la présentation de l'activité scientifique. Ce lexique peut être considéré comme un lexique de genre, n'intégrant pas la terminologie du domaine, mais renvoyant aux concepts mis en œuvre dans l'activité scientifique (*examiner, prouver, réfuter, concluant, hypothèse, examen, encourageant, etc.*) (*ibid.* : 190)

Son corpus est plus élargi que celui de Pecman (*ibid.*) en termes de genres d'écrits scientifiques (articles scientifiques, thèses, rapports, cours) venant des disciplines en sciences humaines et sociales comme la linguistique, l'économie et la médecine. Le corpus d'articles scientifiques est en partie extrait du corpus KIAP⁷ de l'Université de Bergen.

Nous retenons dans ce travail la définition du *lexique scientifique des écrits scientifiques* de Tutin (*opt. cit.*). Nous envisageons à la suite du traitement du corpus de concevoir une liste des expressions polylexicales qui font partie de ce lexique. Cette liste fera par la suite l'objet des analyses syntactico-sémantiques pour être enfin appliquée à des activités pédagogiques. Il s'agit d'une analyse sur un corpus des disciplines en sciences sociales et humaines. Pourtant, nous souhaitons pouvoir proposer une méthodologie d'aide à la rédaction réadaptable à d'autres disciplines y compris les sciences dures. Blumenthal (2007) et Tutin (2010) affirment avec justesse que les listes des mots simples du lexique transdisciplinaire des écrits scientifiques posent quelques problèmes liés à la décontextualisation des mots et à la polysémie des lexèmes. Dans le contexte d'enseignement et d'apprentissage, les simples listes n'ont pas de pertinence pédagogique en soi. Il faudrait donc utiliser ces inventaires lexicaux à bon escient et les recontextualiser. Nous sommes du même avis que Tutin (2010) concernant l'exploitation de ces listes de vocabulaire. Pour rendre ces listes plus exploitables, elles doivent « intégrer des informations sur les fonctions sémantiques et rhétoriques » (*ibid.* : 154). En effet, nous ne préconisons pas d'enseigner la liste de mots aux apprenants. Il est indispensable d'enseigner les unités lexicales dans une chaîne d'écriture avec leurs différentes fonctions rhétoriques. C'est ce que nous envisageons de mettre en place dans le chapitre 6 qui porte sur les réflexions didactiques pour une exploitation de la typologie des unités lexicales.

1.3. Phraséologie transdisciplinaire scientifique : typologie

Il existe actuellement différentes classifications du lexique transdisciplinaire des écrits scientifiques proposées par quelques auteurs, en anglais et en français, en vue d'une application lexicographique à visée didactique. Dans le monde anglophone, plusieurs listes

⁷ KIAP *Kulturell Identitet i Akademisk Prosa*, site de KIAP : <http://kiap.aksis.uib.no/>

ont vu le jour pour des objectifs scientifiques tels que : *Academic Formulas List* (AFL) de Simpson-Vlach & Ellis (2010), qui a pour objectif de construire une liste des segments répétés (ou *formulaic sequences*) pour l'anglais pour objectifs scientifiques et universitaires (par exemple : *on the other, in the first, it is possible* etc.) ; ou bien *Phrasal Expressions List* de Martinez & Schmitt (2012) qui recense 505 expressions polylexicales les plus fréquentes en anglais (*of course, going to, such a* etc.). Même si ces listes de vocabulaire sont conçues à des fins très variées, pour la communication orale (Gougenheim *et al.*, 1964) ou pour constituer le vocabulaire scientifique de base (Phal, 1971 ; Coxhead, 2002), la motivation qui a conduit à leur réalisation reste la même : il s'agit des listes conçues pour des applications didactiques. Avant de décrire notre typologie des séquences polylexicales, nous voudrions passer en revue différentes classifications existantes qui ont inspiré la conception de notre modèle.

1.3.1. Modèle de Pecman

Pecman (2004) mène une analyse contrastive anglais – français des unités phraséologiques⁸ et propose à l'issue de sa recherche un dictionnaire phraséologique composé des unités phraséologiques propres à la langue scientifique. Les objectifs de recherche relèvent à la fois de la linguistique et de la didactique. La recherche vise à proposer des réponses à de nombreuses questions concernant l'identification et les spécificités de la phraséologie de la *Langue Scientifique Générale* (LSG), pour ensuite mettre en place une méthodologie et des matériels linguistiques dans la perspective de création d'un outil d'aide à la rédaction.

Pecman travaille sur un corpus aligné (français – anglais), de sciences dures de 82 800 de mots issus de trois disciplines (la chimie, la physique, la biologie) et composé de 84 textes scientifiques de différents genres (articles, résumés d'articles, rapports d'activité, actes de colloque et comptes rendus). Chaque texte existe en deux versions en anglais et en français. La taille du corpus facilite une exploitation manuelle des données. En voulant créer un dictionnaire phraséologique anglais-français, une unité de traduction consiste en une unité phraséologique anglaise à laquelle se trouve associé un de ses équivalents en français. L'observation et le traitement des unités phraséologiques vont de l'anglais, qui joue le rôle de langue source, au français qui joue donc le rôle de langue cible. Cette direction langue source

⁸ Par « unité phraséologique », Pecman (2004) entend « toute association arbitraire de deux ou plusieurs mots susceptible d'être réutilisée à droit au statut d'unité phraséologiques » (Pecman, 2004 : 28)

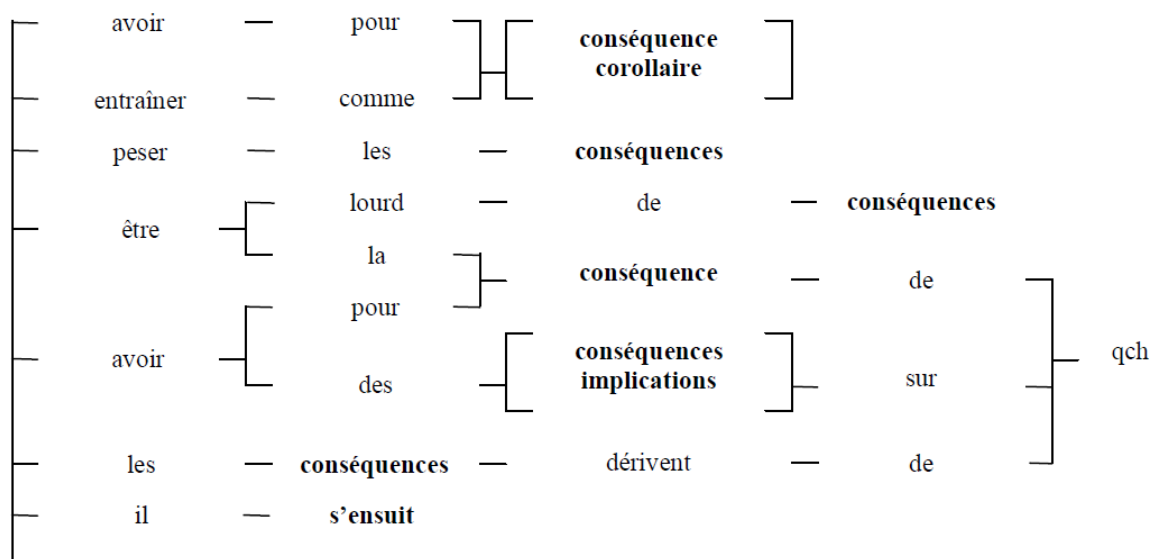
– langue cible est renversée dans la perspective didactique d'aide à la rédaction aux scientifiques francophones.

Pecman (2004) a apporté une réponse à propos des phénomènes collocatifs dans la perspective bilingue, à commencer par la constitution du corpus, la localisation des unités phraséologiques en passant par une mise en place d'un système des descripteurs pour rendre la base de données réutilisable. Elle a proposé d'analyser des données phraséologiques bilingues par deux approches : une approche onomasiologique et une approche syntactico-sémantique. D'une part, l'approche onomasiologique permet de relever des concepts, des thèmes qui structurent la LSG. Chaque concept est associé à plusieurs expressions dont le contenu sémantique est proche de celui-ci. D'autre part, l'approche syntactico-sémantique vise à prendre en compte les propriétés sémantiques et formelles des unités phraséologiques, ainsi que les différentes structures collocationnelles de chaque thème.

Son travail de recherche a permis de déboucher sur un thésaurus de la LSG qui se compose de 125 concepts spécifiques ou hyponymiques relevant de quatre grandes sphères conceptuelles : la scientificité, l'universalité, la modalité et la discursivité. Dans sa recherche, la sphère de la *scientificité* occupe la place la plus importante et inclut tous les concepts qui réfèrent à tout ce qui est de nature exclusivement scientifique ou qui prend une valeur spécifique dans le domaine scientifique, qu'il s'agisse d'une action, d'un objet, d'une entité, d'une abstraction, d'une qualité, etc., avec des concepts comme [EXPERIMENTATION], [EVALUATION], [OBSERVATION] etc. La sphère de l'*universalité* englobe les concepts qui renvoient aux connaissances générales du fait de leur rapport avec la réalité extra-linguistique, par exemple les concepts de [TEMPORALITE], [QUALITE], [CONSEQUENCE], [MANIERE] etc. Les concepts de la *modalité* servent à exprimer un avis, un jugement, une appréciation de la part de l'auteur à l'égard du contenu de son discours, avec comme concepts [PROBABILITE], [NECESSITE]. Dans la sphère de la *discursivité*, on retrouve des unités phraséologiques qui servent à la progression logique du discours et permettent à l'auteur de structurer le contenu de son discours en fonction des objectifs communicatifs qu'il poursuit, avec par exemple des marqueurs exprimant des notions [ADJONCTION] et [CITATION], comme *présentation*, *exemple*, *transition*, etc. Cette ontologie des concepts dans la LSG nous paraît très riche et intéressante sur le plan didactique en proposant à l'apprenant un accès aux collocations par la base ainsi qu'à partir d'un accès onomasiologique. Par exemple, il est possible d'accéder aux expressions liées à la notion [NECESSITE] par le concept de *modalité* auquel appartient également la notion [PROBABILITE]. Néanmoins, certains sous-groupes restent, selon nous, redondants et un peu abstraits en ce que Pecman ne fait pas la distinction

entre la notion et la fonction des concepts, comme le cas des expressions de causalité, de conséquence, qui ont été décrites comme valeurs factuelles et classées dans la sphère de l'universalité, tandis que les mêmes expressions jouent le rôle de connecteurs dans la sphère de discursivité.

Pecman a proposé de mettre en évidence la structure complexe des unités polylexicales selon l'axe syntagmatique et l'axe paradigmatique par l'analyse en schémas collocationnels, modèle emprunté à Altenberg (1998). Cette analyse s'avère efficace pour le type d'unités phraséologiques dans les écrits scientifiques. Comme les unités phraséologiques sont davantage caractérisées par le phénomène de variation que par le phénomène de figement, les schémas collocationnels permettent d'aborder les unités phraséologiques en fonction de la variation des éléments composant l'unité. L'analyse en schémas collocationnels se réalise par l'analyse syntactico-sémantique et consiste à mettre en évidence l'enchaînement des unités linguistiques à l'intérieur d'un domaine notionnel. Les concepts caractéristiques de la LSG et localisés à l'issue de la dernière étape sont soumis à l'analyse en fonction de leur appartenance à une sphère déterminée par des schémas graphiques monolingues et des schémas linéaires bilingues. L'analyse syntactico-sémantique nous semble pertinente au point de vue didactique en mettant à disposition des apprenants des possibilités combinatoires des unités linguistiques sur les plans syntagmatique et paradigmatique, ainsi que les différentes réalisations verbales, nominales, adverbiales, etc. partageant le même domaine notionnel. Un exemple de schéma collocationnel pour exprimer la notion |CONSEQUENCE/ se trouve à la Figure 1-1.



**Figure 1-1 : Extrait du schéma collocationnel pour la notion [CONSEQUENCE]
(D'après Pecman, 2004 : 394)**

Comme l'a montré Tutin (2007 b), certes l'approche notionnelle proposée par Pecman est très intéressante, mais elle pourrait être affinée pour un accès plus ciblé et plus direct aux unités phraséologiques à la fois par la base, mais aussi par le collocatif. À l'instar de Tutin (2007 b), nous ne nions pas l'intérêt des classements fonctionnels et notionnels des unités phraséologiques dans la création d'un dictionnaire de la langue scientifique générale. Néanmoins, un accès trop complexe aux unités lexicales risque selon nous d'augmenter les surcharges cognitives des apprenants, car ceux-ci devraient déterminer à quelle sphère ou à quel concept se trouve rattaché le mot recherché. La difficulté est d'autant plus importante que la frontière entre les sphères n'est pas toujours tranchée, par exemple [AVANTAGE], [DESAVANTAGE], [ORGANISATION], [LIMITE], [ATTENTION], ce qui a été soulevé par Pecman elle-même. En bref, la méthodologie élaborée contribue effectivement à la conception d'un outil rédactionnel qui proposerait un accès flexible aux ressources lexicales moyennant une recherche alphabétique des unités phraséologiques et une recherche sémantique. Ces schémas collocationnels nous inspirent dans la perspective d'aide à la rédaction en ce qu'ils peuvent ouvrir des possibilités dans l'enseignement systématique des paradigmes combinatoires des unités phraséologiques dans une langue étrangère.

1.3.2. Modèle de Riabtseva

Riabtseva (1999) s'intéresse également à l'enseignement des unités phraséologiques en anglais et a conçu un dictionnaire des combinaisons à l'usage scientifique afin d'améliorer les compétences écrites des étudiants et des scientifiques ayant l'anglais comme langue étrangère. Riabtseva travaille sur le métadiscours et les valeurs rhétoriques de ces unités dans les écrits scientifiques en anglais. Elle insiste sur l'importance de créer un support linguistique susceptible d'aider les étudiants à exprimer leurs idées sans passer par la traduction systématique de la langue maternelle. Ce dictionnaire cherche à développer le vocabulaire actif des apprenants en mettant à leur disposition un large choix d'expressions, des combinaisons des mots, des expressions idiomatiques, ainsi que des exemples de l'anglais scientifique.

Riabtseva (*ibid.*) a à la fois une approche empirique et une approche heuristique en partant de genres de textes scientifiques variés en passant par le processus de vérification dans de nombreux dictionnaires. Les utilisateurs peuvent accéder aux expressions soit par une entrée rhétorique ; soit par entrée alphabétique.

D'une part, l'entrée rhétorique est considérée comme un guide qui indique à l'apprenant d'une manière systématique et concise comment organiser les idées dans un article scientifique. Ce guide se compose de trois parties, y compris l'organisation du texte, la modalité (ou *rhetorical devices*) et les commentaires. Tout d'abord, l'apprenant dispose de différents moyens linguistiques, à savoir des séquences polylexicales et des phrases, pour énoncer un problème, définir un matériel ou une méthodologie de la recherche, etc., par exemple : *Now/ Here it is appropriate/ convenient/ desirable/ essential/ fruitful to start with P ; These instances/phenomena tend to form chains/have in common the fact that P*. Ensuite, l'utilisateur peut exprimer son attitude, évaluer et développer une idée par des modèles rhétoriques comme les emphatiques, les auxiliaires (conjonctions, pronoms, etc.), *There are no perfectly acceptable/primarily conventional (forma/functional) ; This research is basically descriptive/ comparative*. Enfin, de nombreux exemples de résumés, de conclusions, d'évaluations, etc. sont systématisés afin d'apporter aux apprenants des indications nécessaires pour la rédaction, *Some researchers suspect/agree (operate on the assumption) that P ; Such analysis/ idea/ view must be rejected in favor of P*. La liste des séquences et des modèles-phrases proposée par Riabtseva est très riche et nous paraît intéressante en ce qu'elle permet aux apprenants d'accéder aux moyens linguistiques par le biais des fonctions rhétoriques. Une telle approche sera très utile pour les activités de production écrite. De plus,

sur le plan architectural, elle aidera les apprenants à acquérir la structure d'un texte scientifique. Néanmoins, une modélisation syntaxique et sémantique de ces groupes de mots permettrait à notre avis de faciliter leur compréhension et leur acquisition chez les apprenants.

D'autre part, l'entrée alphabétique consiste en un dictionnaire de 5 000 mots-clés classés par ordre alphabétique et recense au total 30 000 combinaisons de mots, des expressions et des phrases afin de permettre aux apprenants de se familiariser à l'usage académique des écrits scientifiques. Les termes, les illustrations de ce dictionnaire se retrouvent dans plusieurs types et genres d'écrits scientifiques, et également dans différentes disciplines. Le dictionnaire traite des termes scientifiques traditionnels (*theory, approach, method* etc.), mais aussi des séquences qui expriment la modalité (*can, must, need* etc.), l'évaluation (*reasonable, signifiant, etc.*), la métatextualité (*emphasize, stress, etc.*). En bref, l'approche rhétorique de Riabtseva nous semble intéressante dans la conception des activités pédagogiques pour aider les apprenants à mieux se familiariser avec le lexique de base de la langue scientifique et également avec la structure relativement stabilisée des écrits scientifiques. Néanmoins, nous regrettons quelque peu que les critères de sélection des unités polylexicales ne soient pas précisés. À notre avis, l'organisation des unités polylexicales d'une manière horizontale surcharge la lecture et donne l'impression d'une liste de mots.

1.3.3. Modèle de Granger et Paquot

S'inscrivant dans une optique fonctionnelle, le *Louvain English Academic Purposes Dictionary* (LEAD)⁹ (Granger & Paquot, 2010) nous intéresse particulièrement sur le plan méthodologique. En effet, Granger & Paquot (*ibid.*) ont eu pour objectif de créer un outil d'aide à la rédaction susceptible d'aider les étudiants en mettant à leur disposition des moyens linguistiques de base. Ce dictionnaire se compose de 900 mots et expressions relevés dans un grand corpus de textes académiques (dont le *British National Corpus* et dans le corpus de différentes disciplines) ainsi que du corpus d'apprenants (EFL¹⁰) dont la langue maternelle est, pour une grande majorité, le français. Le LEAD offre aux utilisateurs une riche description du lexique transdisciplinaire des écrits scientifiques en anglais et porte essentiellement sur la phraséologie, en particulier les collocations et les segments répétés.

Sur le plan didactique, le dictionnaire est d'autant plus intéressant qu'il présente les avantages d'un dictionnaire électronique en proposant aux lecteurs différents modes d'accès

⁹ Louvain English Academic Purposes Dictionary LEAD : <http://www.uclouvain.be/en-322619.html>

¹⁰ English as a Foreign Language

comme le *mode sémasiologique* (de la forme au sens) et le *mode onomasiologique* (du sens à la forme) par le biais d'une liste des fonctions rhétoriques ou organisationnelles du discours scientifique. D'une autre part, le *Louvain EAP Dictionary* est également un *dictionnaire semi-bilingue* en ce sens que l'utilisateur a la possibilité de choisir sa langue maternelle et faire des recherches par des entrées lexicales traduites dans cette langue. Le choix de la langue maternelle des apprenants joue un rôle important dans l'identification des erreurs et des problèmes spécifiques de chaque langue. Il sert aussi à analyser les influences sur la traduction de la langue maternelle à la langue cible. Toutes les erreurs rencontrées par les apprenants sont enregistrées et signalées par la suite aux apprenants si ceux-ci choisissent la même langue maternelle, par exemple, l'erreur de traduction de *prétendre* en français pour le mot *pretend* en anglais par exemple. De plus, l'apprenant peut choisir le corpus sur lequel il veut travailler. Il existe actuellement trois disciplines couvertes : la gestion, la médecine, la linguistique. Cette étape constitue une préparation interactive où l'outil lexicographique aide les utilisateurs à identifier et à spécifier leurs besoins concrets avant d'être guidés vers les bases de données correspondantes.

Outre le mode sémasiologique qui permet aux apprenants de rechercher un mot par son lemme, ils peuvent également accéder à de nombreuses unités phraséologiques de même concept par le mode onomasiologique. L'*accès onomasiologique* s'effectue par le biais d'une liste des 18 fonctions rhétoriques et organisationnelles, particulièrement représentées dans le discours scientifique. On y retrouve des concepts pour introduire une thématique : énumérer, ajouter une information, ou exprimer la causalité, donner un exemple, reformuler, etc. Pour chaque fonction, l'utilisateur peut accéder à une liste des éléments lexicaux typiques dans le discours scientifique et pour chaque catégorie (noms, verbes, adjectifs, adverbes, conjonctions et prépositions) (Figure 1-2). Cette classification nous paraît très intéressante dans la perspective de l'aide à la rédaction. Les approches menées par les auteurs nous inspirent beaucoup dans la conception de notre typologie des séquences polylexicales, car nous envisageons nous aussi d'adapter l'approche onomasiologique dans l'enseignement et l'apprentissage du lexique transdisciplinaire des écrits scientifiques comme moyen susceptible d'aider les étudiants à développer leurs compétences lexicales et phraséologiques. Nous y reviendrons plus en détail dans le chapitre 6.

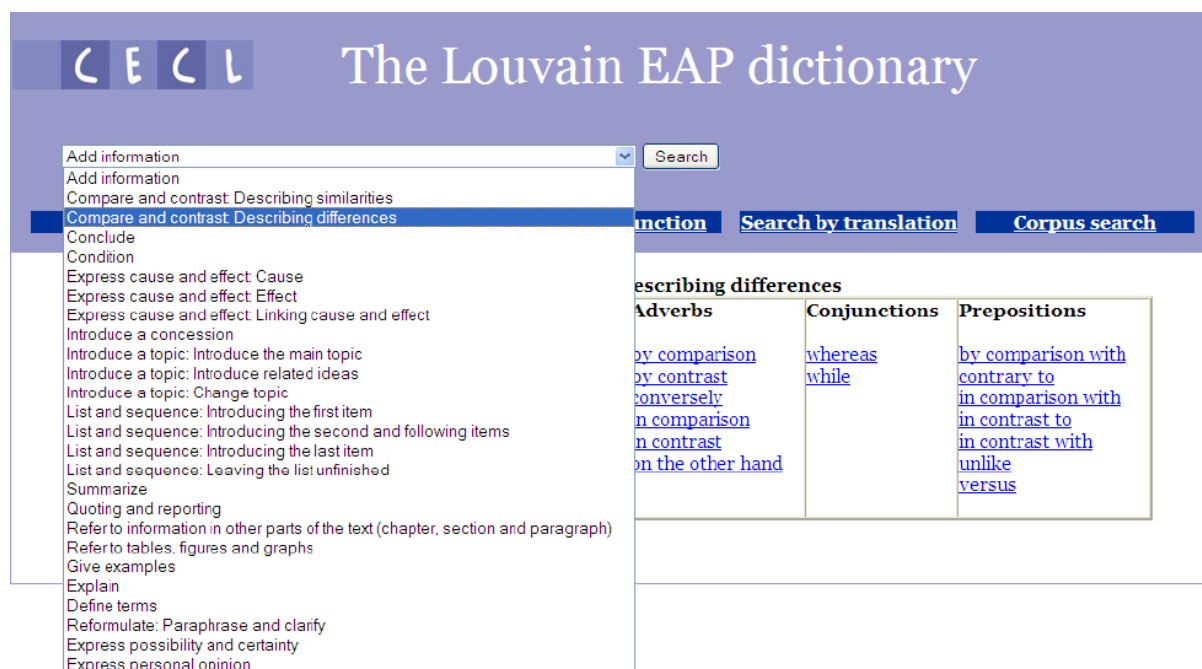


Figure 1-2 : Accès onomasiologique dans LEAD (d'après Granger et Paquot, 2010 : 323)

En bref, les trois typologies que nous avons présentées dans cette partie ont chacune une vision didactique différente, mais elles partagent le même objectif d'aide à la rédaction scientifique. Dans notre travail, nous envisageons de tirer profit de l'approche onomasiologique en proposant une entrée par les fonctions discursives et rhétoriques. Une telle approche nous apporte de nombreux avantages que nous allons expliciter dans le chapitre 4 et 6.

1.4. Intérêt pour le lexique transdisciplinaire scientifique

Afin de pouvoir proposer des approches adaptables à l'enseignement de ce lexique, il nous semble indispensable de détecter avant tout les problèmes que rencontrent les apprenants non natifs dans la rédaction de leurs écrits (mémoires, rapport, etc.) en français. C'est ce que nous allons expliciter dans cette partie et nous soulignerons par la suite l'importance de l'enseignement du lexique transdisciplinaire scientifique.

Comme nous l'avons vu antérieurement, le lexique transdisciplinaire dans les écrits scientifiques ne concerne pas uniquement les mots isolés, mais également les combinaisons de mots dont nous pouvons distinguer principalement deux sortes, à savoir les collocations et les séquences récurrentes. Le premier groupe, les collocations, ont été définies par Tutin & Grossmann (2002) comme des expressions binaires, constituées de deux éléments

linguistiques, apparaissant fréquemment en cooccurrence et entretenant une relation syntaxique ; un élément, la base, garde son sens habituel ; l'autre, le collocatif est moins prédictible et dépend du premier, par exemple : *valider une hypothèse, rôle crucial*, etc. Le deuxième groupe englobe pêle-mêle les séquences récurrentes ou des suites de mots contigus dont l'appellation est très variée, par exemple : *lexical bundles* (blocs lexicaux) (Biber, Conrad & Cortes, 2004 a), *formulaic sequences* (formules du discours) (Wray, 2002), *sequences of words* (séquences de mots) (De Cock, 2004), par exemple : *il en est de même, en revanche*, etc.

Plusieurs études ont été menées avec l'objectif de comprendre comment les apprenants non natifs utilisent ces éléments phraséologiques dans leurs productions écrites. Elles relèvent que les erreurs les plus fréquentes restent la combinatoire lexicale (Simard, 1994) qui cache parfois d'autres problèmes au plan syntaxique et sémantique (Paquot, 2010 ; Osborne, 1994 ; De Cock, 2004).

Au plan syntaxique, Nesselhauf (2005) a relevé un taux d'erreurs de 17 % pour les séquences libres contre plus de 25 % des erreurs des collocations. Pour les collocations, l'erreur la plus fréquente consiste en un mauvais choix des verbes et elle concerne 56 % des étudiants allemands en anglais langue étrangère. Osborne (2008) précise trois erreurs principales dans la production écrite des étudiants non natifs de l'anglais qu'il nomme *blending*, *bonding*, et *burying*. Par *blending*, l'auteur entend l'association des éléments pour former une unité plus grande qui n'est pourtant pas conforme aux règles grammaticales. Par *bonding*, Osborne évoque un autre problème souvent rencontré chez les étudiants. Il s'agit de l'insertion au milieu d'une collocation ou une association un autre élément, ce qui produit des associations inappropriées, par exemple : *follow blindly everything* etc. Par *burying*, la collocation perd sa valeur grammaticale et devient moins saillante quand elle est intégrée par une unité plus large. L'auteur donne l'exemple de *love tender*, et l'ajout de la phrase complète au milieu de la collocation lui fait perdre en quelque sorte le sens du verbe *love*. Ces erreurs peuvent être expliquées, selon nous, par un manque de connaissances du phénomène des collocations de la part des apprenants.

Au plan sémantique, en analysant les écrits des étudiants d'anglais langue étrangère dans son étude, Paquot (2010) constate que ceux-ci rencontrent des difficultés concernant les valeurs sémantiques de certains moyens cohésifs (*cohesive devices*), ce qui révèle l'importance d'aborder les valeurs sémantiques de ces unités phraséologiques dans l'enseignement. En outre, Paquot (*ibid.*) remarque une redondance dans les moyens cohésifs ou l'utilisation des valeurs sémantiques incompatibles. Quant à la position des connecteurs

dans les écrits des étudiants non natifs, ceux-ci semblent privilégier la position en tête de phrase, la position médiane n'étant pas utilisée, contrairement aux écrits des natifs. Ces erreurs d'ordre sémantique sont en partie dues selon nous à un enseignement insuffisant de ces éléments. C'est pourquoi nous envisageons de mettre en place des activités didactiques qui permettent de faire distinguer aux apprenants différentes valeurs sémantiques des MD. Ces remarques nous intéressent et nous apportent des éléments pertinents à prendre en compte dans l'analyse linguistique que nous allons effectuer dans le chapitre 5. Nous allons traiter à la fois les propriétés syntaxiques et sémantiques des MD, c'est-à-dire leur position et leurs valeurs sémantiques. Nous effectuerons dans le chapitre 6 une analyse de corpus de mémoires d'étudiants vietnamophones pour relever les erreurs les plus fréquentes dans l'utilisation de ces unités lexicales.

En particulier, De Cock (2004) constate un manque de *routinized ways* (routines argumentatives) pour établir l'interaction et le rapport entre les interlocuteurs. Paquot (2010), De Cock (2004) remarquent une sur-utilisation des certaines collocations les plus fréquentes en anglais, mais une sous-utilisation des unités polylexicales dans les écrits des étudiants (en anglais langue étrangère).

Les raisons de ces erreurs sont nombreuses : l'influence de la langue maternelle (Granger & Paquot, 2010), le répertoire lexical restreint d'où la sur-utilisation de certains mots ou groupes de mots, le manque de connaissances sur le registre (Paquot, 2010), etc. D'après ces enquêtes, les erreurs lexicales sont peu prises en compte par les étudiants dans la révision de leur texte au détriment d'autres dimensions grammaticales et orthographiques, ce qui entraîne par la suite un manque de cohérence dans les productions écrites des étudiants même au niveau avancé (Paquot, 2010).

Anctil (2006) a mené une étude auprès des étudiants qui sont inscrits au cours *français écrit pour futurs enseignants* et qui ont le français comme langue maternelle. Il a montré que les problèmes liés aux propriétés de combinatoire sont les problèmes les plus fréquents (30 %) chez les étudiants. Ces difficultés observées chez les apprenants s'expliqueraient par le manque d'information « explicite et systématique dans les ouvrages de référence » (Anctil, 2005 : 109) et aussi par le manque de formation de la part des enseignants. Les enseignants de français associent souvent des erreurs dites de vocabulaire au mauvais choix de mots ou les catégorisent comme des maladresses d'expression. L'enseignant de FLE qui doit enseigner à un public scientifique se trouve confronté à un problème difficile, d'autant plus que les méthodes de français général ne répondent que partiellement aux besoins langagiers de ces publics.

Il est nécessaire donc de mettre l'accent sur l'importance d'un enseignement/apprentissage plus systématique de ces unités phraséologiques, essentiellement des séquences récurrentes (Howarth, 1998 ; Gledhill, 2000 ; De Cock, 2004). Biber *et al.* (2004) les considèrent comme « les blocs de base du discours ». Selon Howarth (1998), l'enseignement de ces moyens lexicaux est susceptible d'aider les apprenants à répondre aux « normes stylistiques d'un genre particulier » qui consiste non seulement en « les choix grammaticaux ou lexicaux, mais aussi la sélection des unités polylexicales d'une manière appropriée » (*ibid.* : 186). En bref, un enseignement de ces unités phraséologiques sera bénéfique pour les étudiants 1/ pour leur donner plus d'aisance dans l'écriture ; 2/ pour qu'ils s'intègrent mieux dans une communauté par l'acquisition des normes stylistiques pour un registre particulier (Wray, 2002).

Conclusion

Dans cette partie, nous avons délimité le domaine de notre recherche et défini le lexique qui nous intéresse, c'est-à-dire le lexique transdisciplinaire dans les écrits scientifiques. En raison des spécificités de ce lexique, les apprenants rencontrent souvent des erreurs d'ordre syntaxique, sémantique et organisationnel. Les raisons sont multiples, mais la raison principale est relative au manque d'enseignement systématique de ces éléments phraséologiques malgré leur importance dans le discours. Nous souhaitons proposer dans ce travail un traitement des séquences lexicalisées à fonction discursive dans les écrits scientifiques pour une application didactique du FLE. Passons maintenant au chapitre 2 où nous allons définir plus en détail notre objet d'études, les MD ou les séquences lexicalisées à fonction discursive.

Chapitre 2.

Marqueurs discursifs

Il est à préciser que notre objet d'étude s'intègre à la fois dans le domaine de la phraséologie dans une perspective étendue, comme mentionné par Legallois & Tutin (2013), mais nous nous intéressons aux fonctions discursives de certaines suites de mots récurrentes dans les écrits scientifiques, que nous allons appeler les « séquences lexicalisées à fonction discursive » ou les « MD polylexicaux ».

Commençons donc par situer notre objet de recherche. Parmi différentes typologies de la phraséologie en langue générale, celle de Burger (1998) nous intéresse particulièrement, car elle prend appui sur la fonction qu'occupent les unités phraséologiques dans le discours. Burger (*ibid.*) distingue trois catégories : les unités phraséologiques (UP) à fonction référentielle, celles à fonction structurelle et celles à fonction communicative. Premièrement, les UP à fonction référentielle comprennent les idiomes, les collocations, les proverbes, les séquences idiomatiques qui renvoient aux objets, aux référents et à la réalité. Ensuite, les UP à fonction communicative ainsi que les routines conversationnelles sont utilisées comme des moyens pour amorcer, maintenir et clore la conversation (*bonjour, je vais bien*, etc.). Enfin, les UP à fonction structurelle jouent le rôle d'organisateur du texte. Néanmoins, il est dommage que cette dernière catégorie ne soit pas très développée dans les travaux de Burger, car elle ne constitue qu'un petit groupe des UP.

Tutin (2014) a adapté cette classification de Burger (*op.cit.*) au lexique transdisciplinaire scientifique, tout en s'appuyant sur des critères fonctionnels. Elle a proposé un classement plus approfondi des séquences polylexicales (SP) dans les écrits scientifiques : a/ les SP à fonction référentielle (*jouer un rôle, objet de recherche*, etc.), b/ les SP à fonction discursive ou appelées marqueurs de discours (*en revanche, par ailleurs*), c/ les SP à fonction interpersonnelle (*il est souhaitable, contre toute attente*) et d/les routines sémantico-

sémantiques (*nous reprenons le modèle développé par X, etc.*). À la différence de Burger (*op.cit.*), Tutin porte un intérêt particulier aux marqueurs de discours et rejoint à Siepmann (2007) en affirmant que ces expressions sont assez mal recensées et étudiées, en particulier les « marqueurs de deuxième niveau », c'est-à-dire les marqueurs de fréquence moyenne. Cette typologie fonctionnelle correspond à nos objectifs linguistiques et didactiques, car nous envisageons d'analyser les séquences polylexicales à fonction discursive sur le plan syntaxique et sémantique afin de proposer par la suite une approche d'enseignement et d'apprentissage de ces unités linguistiques dans la perspective d'une aide à la rédaction.

Nous allons maintenant définir les MD et leurs propriétés. Ensuite, nous allons mentionner quelques difficultés rencontrées par les apprenants lors de la rédaction de leur travail de recherche. Nous allons effectuer un petit état de l'art par la suite pour comprendre comment les MD sont introduits dans les manuels et les grammaires de FLE. Enfin, nous décrirons quelques typologies des MD qui nous ont inspirée dans la conception de notre typologie des séquences lexicalisées à fonction discursive.

2.1. Intérêt pour les marqueurs discursifs

2.1.1. Définition

La définition des MD n'est pas une chose aisée. Ces unités linguistiques très diverses portent différentes appellations selon les écoles et les théories, par exemple : *ponctuants*, ou *ponctuation (discursive et métadiscursive)* (dans une perspective prosodique), ou *connecteurs logiques*, *particules (discursives, énonciatives)* (Fernandez, 1994), *marqueurs de glose* (Steuckardt & Niklas-Salminen, 2005). Ces unités sont souvent appelées *connecteurs logiques* dans de nombreuses méthodes de FLE. Depuis quelques années, on entend de plus en plus parler de *marqueurs discursifs* ou *discursive markers* en anglais. De nombreuses études considèrent ces marqueurs comme des moyens de repérer dans les discours les relations de cohérence. Selon Laignelet (2004), les marqueurs de discours « peuvent être envisagés comme des marques linguistiques non référentielles méta-discursives et dont le rôle est de signaler de façon plus ou moins explicite l'organisation du texte » (*ibid.* : 20).

Dans cette recherche, nous nous intéressons tout particulièrement aux MD polylexicaux, pour deux raisons. D'une part, nous supposons que les apprenants sont plus habitués aux MD monolexicaux, car les manuels de FLE les traitent davantage que les MD polylexicaux. D'autre part, nous supposons que les MD polylexicaux sont parfois plus

complexes que les MD monolexicaux en raison de leur structure syntaxique (*par contre, d'autant plus*, etc.)

Comme le terme *marqueurs discursifs* peut comprendre à la fois des marqueurs du discours oral et des monolexicaux (*tiens, donc, bien, etc.*), nous utilisons une notion plus précise de notre point de vue : les *séquences lexicalisées à fonction discursive*. Par *séquences lexicalisées*, nous entendons une suite de mots fonctionnant comme une unité lexicale. Il peut s'agir de locutions adverbiales (*pour commencer*), prépositionnelles (*suite à*), ou conjonctives (*de façon à ce que, pour que*). Une locution adverbiale ou prépositionnelle dispose du même fonctionnement dans la phrase qu'un adverbe ou une préposition. Le terme *lexicalisé* renvoie au mécanisme de formation de nouvelles formes composées qui permet d'enrichir la classe grammaticale en question.

Nous définissons les séquences lexicalisées à fonction discursive comme des éléments linguistiques qui établissent les relations entre les segments textuels, c'est-à-dire les propositions et les séquences qui composent le texte. Ces éléments indiquent aussi les articulations du discours ; ils assurent l'enchaînement entre les propositions et la structuration hiérarchisée du texte en ensembles de propositions. Ils servent donc à organiser le texte et à structurer le discours, par exemple *pour commencer, par ailleurs, en d'autres termes, en premier lieu, par conséquent*, etc.

Les MD consistent en des moyens nécessaires à la structuration du discours, mais constituent souvent des difficultés dans le développement des compétences rédactionnelles chez les étudiants étrangers. Comme nous l'avons précisé antérieurement, notre sujet de recherche est étroitement lié aux besoins linguistiques des étudiants étrangers, en particulier dans la rédaction de leur rapport et leur mémoire de fin d'études. Nous partons du constat de plusieurs auteurs ayant étudié la cohérence dans les productions écrites d'apprenants avancés, parmi lesquels Osborne (1994) qui remarque une « maîtrise imparfaite des phénomènes de cohésion » (*ibid.* : 205) dans ces écrits. Cette maîtrise insuffisante de la cohérence à l'écrit est souvent due à l'absence ou à l'abus des organisateurs textuels, ce qui renvoie au lecteur l'image d'une accumulation désorganisée des idées au lieu d'un objet bien construit.

Nous considérons les MD en général et les séquences lexicalisées à fonction discursive en particulier comme un des moyens susceptibles de sensibiliser les apprenants à la cohérence dans le texte. Elles font partie des moyens lexicaux pour les aider à mieux structurer leur écrit. Un enseignement des séquences lexicalisées à fonction discursive sera très bénéfique pour les activités de compréhension écrite et de production écrite. En effet, les MD, par leurs fonctions de guidage de lecteur et de structuration du texte (comme nous allons l'aborder dans le

chapitre 4), serviront à développer chez les apprenants une compétence discursive et à faciliter la lecture par le repérage des structures discursives. Nous sommes du même avis que Allouche & Maurer (2011) quand ils insistent sur l'importance de l'enseignement aux étudiants des moyens lexicaux pour créer la cohérence textuelle : « par delà les connaissances relatives aux contenus des textes et leur organisation structurelle, un entraînement à la maîtrise des outils linguistiques servant à construire la cohérence d'un texte est nécessaire » (*ibid.* : 36). C'est pourquoi l'enseignement des MD permet aux apprenants d'« anticiper l'information » et de repérer « la distribution hiérarchisée des éléments de contenu selon un ordre donné » (*ibid.* : 59).

De plus, en travaillant sur les écrits scientifiques, les apprenants ont besoin d'être mieux guidés dans la construction de leur discours par des modèles textuels authentiques. Nous insistons donc sur l'importance d'intégrer les séquences lexicalisées à fonction discursive dans les cursus de formation au service du public scientifique pour qu'elles deviennent un objet d'enseignement et d'apprentissage plus systématique.

2.1.2. Propriétés syntaxiques et sémantiques

Avant de vérifier comment les MD sont utilisés en classe de langue, il nous semble important d'étudier leurs différentes propriétés, afin de comprendre les difficultés pour les locuteurs non natifs. Dostie & Pusch (2007) ont fait un bilan sur différentes propriétés de ces éléments linguistiques, aux plans morphologique, syntaxique et sémantique. Sur le plan morphologique, ces marqueurs sont invariables et appartiennent à différentes catégories syntaxiques y compris les adverbes, les conjonctions et les prépositions. Ce caractère hétérogène des MD constitue une difficulté dans la classification de ces unités lexicales, car la majorité des MD se construisent « à partir de la décatégorisation et recatégorisation d'unités appartenant, initialement, à d'autres classes grammaticales » (*ibid.* : 4), par exemple : *vraiment, bon, disons*, etc. ou les locutions adverbiales formées de deux ou plusieurs mots : *par contre, par ailleurs, en fait*, etc. Pourtant, le seul critère morphosyntaxique ne permet pas d'affirmer que telle ou telle forme appartient à la catégorie des MD, comme *forcément, clairement, certainement* sont à la fois des adverbes et des MD tandis que *gentiment, faussement* ne sont que des adverbes.

Sur le plan syntaxique, les MD sont optionnels et peuvent occuper différentes positions dans l'énoncé. À la différence d'autres compléments, les MD n'entrent pas dans la structure argumentale de l'énoncé, comme le complément d'objet direct et le complément

d'objet indirect. Leur présence ou leur absence n'influe pas sur la valeur de vérité des énoncés (Dostie & Pusch, *op.cit.*). Cette caractéristique des marqueurs discursifs coïncide avec celle décrite par Riegel *et al.* (2009) : les connecteurs appartiennent à la catégorie des MD et « ne font pas partie intégrante des propositions et n'y exercent aucune fonction syntaxique, mais assurent leur liaison et organisent leurs relations » (*ibid.* : 1044). En ce qui concerne l'enseignement et l'apprentissage des MD, leur grande mobilité va déstabiliser les étudiants, comme l'a remarqué Delahaie (2011).

Sur le plan sémantique, selon Dostie & Pusch (*op.cit.*), même si les MD ne contribuent pas au contenu propositionnel des énoncés, ils apportent des informations sur l'énonciation et le discours. Il faudrait donc traiter des valeurs sémantiques des MD dans l'enseignement de ces unités lexicales. Le traitement des MD présente une difficulté supplémentaire associée à leur polysémie ou leur polyfonctionnalité (terme suggéré par Paillard & Vu (2012)) en raison : a/ de la variation de leur portée (énoncé, proposition, mot isolé ou syntagme) b/ de leur position dans l'énoncé (initiale, médiane, globale). Paillard & Vu (*ibid.*) définissent la portée comme « la séquence à laquelle le MD confère un statut discursif » (*ibid.* : 21). Nous allons la définir en détail dans le chapitre 5, car la portée fait partie des paramètres les plus importants dans notre analyse des MD (Cf. Chapitre 5). Reprenons l'exemple de *enfin* dans Paillard & Vu (*ibid.*). *Enfin* peut avoir différentes valeurs d'emploi en fonction de sa position et de sa portée. Il peut exprimer une reformulation : *on était arrivé. Enfin, presque*, ou une résignation : *Enfin, bon j'y vais*, ou bien pour exprimer un aboutissement : *Le repos, enfin*.

En bref, ces différentes caractéristiques des MD constituent des difficultés majeures et risquent de déstabiliser les locuteurs non natifs en raison de leur complexité et de leur variété. Même si la position des MD peut être variée, elle n'est pas aléatoire. Un changement de position peut entraîner un changement de sens. Notre objectif sera donc de relever toutes les régularités de fonctionnement de ces unités lexicales, comme nous le verrons avec une sous-catégorie de MD dans le chapitre 5. Nous pouvons résumer les propriétés principales des MD dans le Tableau 2.1.

Propriété	Marqueurs discursifs
Morphosyntaxique	<ul style="list-style-type: none"> - Appartenance à différentes catégories syntaxiques - Invariable
Syntaxique	<ul style="list-style-type: none"> - Optionnel - Positions variables dans l'énoncé
Sémantique	<ul style="list-style-type: none"> - Polyfonctionnalité en raison de la variété de la portée et de la position

Tableau 2.1 : Propriétés principales des MD

2.2. Marqueurs discursifs : enseignement/apprentissage en FLE

Dans cette partie, nous allons relever quelques erreurs fréquentes dans l'utilisation des MD dans les mémoires des étudiants. Il s'agit d'erreurs syntaxiques, sémantiques et discursives. Nous allons par la suite faire un petit état de l'art de l'enseignement/apprentissage de ces éléments dans les manuels de FLE et dans les livres de grammaire de FLE pour sonder comment ces unités lexicales sont abordées en classe de langue.

2.2.1. Utilisation des marqueurs discursifs par des apprenants non natifs

2.2.2.1. Erreurs d'ordre syntaxique

Ayant pour ultime objectif d'améliorer la compétence rédactionnelle des étudiants vietnamiens, Le (2013) a cherché à comprendre l'utilisation des connecteurs argumentatifs dans leurs mémoires de fin d'études et à détecter les erreurs les plus fréquentes afin d'en proposer des solutions et des stratégies d'enseignement. Son corpus de recherche d'une taille de 313 994 mots (titre, introduction, développement, conclusion, notes, annexes) se compose de vingt-et-un mémoires de licence de la Filière francophone de Géographie de l'Agence Universitaire de la Francophonie (AUF) à Ho Chi Minh ville, soutenus entre 2006 et 2010. Par des analyses minutieuses, elle a relevé des erreurs et anomalies d'ordre syntaxique, sémantique et relatives à l'organisation des idées. Sur le plan syntaxique, les erreurs et

anomalies d'utilisation des argumentatifs consistent en des distributions syntaxiques peu fréquentes à l'écrit. Dans les mémoires, beaucoup de connecteurs sont utilisés en début de phrase comme *alors*, *donc*, tandis que selon les normes d'usage de ces connecteurs, l'antéposition n'est pas permise ou bien elle ne peut être mobilisée de façon fréquente et systématique. Cette anomalie est due, d'une part, à l'influence de la langue maternelle (le vietnamien), et d'autre part, à l'enseignement insuffisant de ces connecteurs. En fait, en langue vietnamienne, les connecteurs s'utilisent en tête d'énoncé. En outre, Le (*ibid.*) a remarqué que dans les programmes de français général et dans les programmes de Français sur Objectif Spécifique (FOS), les connecteurs sont présentés en abondance dans les dialogues et les discours oraux, alors que les connecteurs sont très peu introduits dans les contextes écrits.

Les erreurs concernent la transgression des règles d'usage du connecteur, soit la syntaxe du connecteur mobilisé, soit la syntaxe de la proposition où est intégré le connecteur. Par exemple, dans leurs mémoires, les étudiants ont utilisé *parce qu'* après un groupe nominal ou un infinitif au lieu d'une proposition, ou bien ce connecteur causal est utilisé en combinaison avec un autre connecteur consécutif comme l'association *par conséquent... donc* ou *parce que... alors*. Toutes ces erreurs peuvent s'expliquer par l'influence du vietnamien où la relation de cause à conséquence s'effectue par une paire de connecteurs causal et consécutif *Vi A cho nen B (parce que A alors/donc B)*. Il est à noter que l'utilisation erronée de ces connecteurs manifeste également une non-maîtrise de la règle syntaxique des connecteurs en français par les étudiants.

2.2.2.2. Erreurs d'ordre sémantique

Le (*op.cit*) a constaté des erreurs d'ordre sémantique dans l'emploi inapproprié des connecteurs ou dans l'utilisation excessive de ces connecteurs. Il s'agit de la non-distinction des connecteurs consécutifs sémantiquement proches comme *alors*, *donc*, *aussi*. Comme l'étudiant cherche à marquer des déductions objectives et subjectives, il est amené à faire varier les connecteurs consécutifs dans le discours sans pour autant acquérir les différentes nuances discursives de ces termes. Par exemple, l'utilisation de *donc*, *aussi* à la place d'*en effet*, l'emploi d'*alors* au lieu d'*alors que*. Le (*op. cit.*) a relevé également la sur-utilisation des connecteurs chez certains scripteurs, par exemple l'utilisation dans la même séquence de deux connecteurs *alors... donc*, ou *ainsi... pourtant*. Toutes les utilisations erronées s'expliquent par la non-maîtrise des outils linguistiques utilisés, ce qui pose par la suite des problèmes de cohérence, et de progression thématique.

2.2.2.3. Erreurs d'ordre discursif

Le (*op. cit.*) a repéré des anomalies relatives à l'organisation des idées dans une séquence, que ce soit l'organisation confuse des idées ou bien la fragmentation des idées. La confusion dans l'agencement des idées rend le texte incompréhensible et l'utilisation de certains connecteurs devient dans ce cas inefficace, inappropriée. De plus, l'auteure a constaté des anomalies en ce qui concerne la fragmentation des contenus en plusieurs séquences, l'un des moyens de fragmentation étant par ailleurs le positionnement des connecteurs (*cependant, en plus, par ailleurs*) en tête d'énoncé. Toutes ces anomalies, qui gênent la compréhension de la progression thématique, sont dues à un manque de connaissances de la mise en texte de la cohésion et de la cohérence argumentative en français, dans le discours.

Diep (2008) a effectué une étude de corpus de productions écrites d'étudiants vietnamiens dans le discours agronomique et a observé comment les apprenants s'approprient les phénomènes de cohérence. Il a constaté que, si les apprenants vietnamiens acquièrent bien les « règles grammaticales au niveau de la construction phrastique » (*ibid.* : 364), ils rencontrent des difficultés dans l'organisation du texte, ce qui se manifeste par de nombreuses ruptures dans la continuité thématique, des maladroites conduisant à des progressions illogiques, mais aussi une faible fréquence des connecteurs logiques et textuels (la reformulation, la récapitulation, la justification, etc.). Ces différentes erreurs qu'a relevées Diep (*ibid.*) peuvent s'expliquer par un manque de connaissances des apprenants sur les phénomènes de cohérence. Les apprenants mobilisent les efforts au plan grammatical et lexical au détriment de l'organisation globale du texte.

Une autre étude sur des corpus d'apprenants étrangers a été effectuée par Glorieux *et al.* (2006). Soucieux de la maîtrise de la langue française chez les étudiants étrangers, ils souhaitent élaborer des outils d'aide à la rédaction d'écrits scientifiques. Ils envisagent, d'une part, d'informer les enseignants des difficultés rencontrées par les étudiants et de leur fournir des pistes pédagogiques ; d'autre part, ils cherchent à proposer des pistes méthodologiques et théoriques aux étudiants afin de répondre aux exigences liées aux divers types d'écrits. Pour répondre à ces objectifs, Glorieux *et al.* (*ibid.*) ont mené des études de cas et analysé des productions d'étudiants de niveau B1 du CECRL, venant de différentes facultés de sciences humaines et de sciences appliquées d'universités belges, afin de comprendre les exigences universitaires et repérer les difficultés liées aux différents types de travaux (résumé d'un article scientifique, lecture critique de texte, production de réflexions personnelles, rapport de recherche). Cette analyse nous a apporté des éclairages sur les difficultés et les limites dans

les écrits des étudiants. Les difficultés observées sont à la fois d'ordre linguistique, pragmatique et organisationnel. Plusieurs problèmes liés à la non-maîtrise de la langue ont été relevés, comme l'imprécision du vocabulaire, l'inadéquation du niveau de langue (répétitions, redondances, phrases mal construites, etc.), et également le non-respect des conventions d'écriture du résumé (utilisation d'abréviations, de tirets et de flèches). Un constat qui nous intéresse tout particulièrement est lié à l'insuffisance ou l'absence d'organiseurs textuels et une faiblesse des techniques de reformulation. C'est pourquoi il nous semble important de mettre à la disposition des apprenants des moyens lexicaux qu'ils pourront réinvestir par la suite dans leur écrit. En particulier, il est nécessaire de leur montrer simultanément comment les insérer dans une chaîne de l'écrit et le processus d'écriture. À notre avis, la reformulation fait partie des compétences importantes à acquérir et nous reviendrons sur ce point dans le chapitre 5 qui est consacré à une étude de cas des marqueurs de reformulation.

À l'instar de Glorieux *et al.* (*op. cit.*), qui analysent les difficultés propres à l'activité d'écriture pour un sujet qui ne rédige pas dans sa langue maternelle, Allouche & Maurer (2011) ont constaté un manque de réflexion sur certaines difficultés d'écriture propres aux apprenants de FLE et un manque d'outils concrets en termes d'activités et de séances pour pallier ces difficultés. Parmi les compétences nécessaires à la production d'écrits comme la prise en compte de la scripturalité, le choix du mot juste, les auteurs insistent sur l'entraînement à la maîtrise des outils linguistiques servant à construire la cohérence d'un texte au-delà des connaissances relatives aux contenus des textes et à leur organisation structurelle.

Pour conclure cette partie sur l'utilisation des connecteurs argumentatifs et des organisateurs textuels dans les productions écrites des étudiants non natifs, outre les erreurs liées à l'influence de la langue maternelle, nous notons des erreurs grammaticales et stylistiques révélant des faiblesses chez les scripteurs en langue cible. Les auteurs insistent tous sur l'absence d'un enseignement plus systématique des procédés et techniques argumentatifs, essentiellement adaptable au public scientifique. Il est donc indispensable avant tout de considérer les MD comme un objet d'enseignement pertinent pour les étudiants. Nous pouvons résumer les erreurs les plus fréquentes dans l'utilisation des MD chez les locuteurs non natifs dans le tableau suivant :

Erreurs	Descriptions
Syntaxiques	<ul style="list-style-type: none"> - Des distributions peu fréquentes à l'écrit - Transgression des règles grammaticales, syntaxe défectueuse
Sémantiques	<ul style="list-style-type: none"> - Utilisation inappropriée des connecteurs - Sur-emploi de certains connecteurs
Organisationnelles	<ul style="list-style-type: none"> - Organisation confuse d'idées, fragmentation d'idées. - Absence ou insuffisance des connecteurs, des organisateurs textuels. - Rupture dans la continuité thématique - Manque de compétence de reformulation

Tableau 2.2 : Erreurs fréquentes de l'utilisation des MD chez les locuteurs non natifs

En partant de l'hypothèse selon laquelle le manque de guidage dans les programmes d'enseignement est une des causes de la faiblesse de l'utilisation des organisateurs textuels chez les étudiants, nous allons analyser comment ces unités linguistiques sont traitées dans les méthodes de FLE et dans les livres de grammaire de FLE. Quelles sont les stratégies d'apprentissage ou les activités mises en place pour aider les étudiants non natifs à acquérir ce lexique ?

2.2.2. Enseignement/apprentissage des marqueurs discursifs en FLE

À notre connaissance, les MD ne sont pas absents des travaux linguistiques et y sont traités d'une manière bien précise. On peut citer de nombreuses recherches qui se consacrent à un seul MD, parmi lesquelles l'analyse sur *en fait, de fait* (Danjou-Flaux, 1980), les marqueurs de reformulation paraphrastique (Gülich & Kotschi, 1983), *pour autant, pourtant* (Anscombe, 1983), les opérations de reformulation (Rossari, 1994), etc. Ces descriptions linguistiques sont très intéressantes et importantes sur les plans théorique et méthodologique en ce qu'elles nous apportent une vision très détaillée d'un marqueur. Néanmoins, ces descriptions doivent être didactisées pour être utilisées pour un public FLE en raison de leur complexité scientifique et de leur spécialisation linguistique (tous les apprenants ne sont pas

des linguistes) (Garnier & Savage, 2011). Il s'avère difficile, du point de vue didactique, d'intégrer dans les activités d'enseignement et d'apprentissage toutes les analyses centrées sur une seule unité. Tout cela explique l'importance des recherches qui portent sur la transposition didactique à partir des analyses linguistiques existantes. Pourtant, il en existe encore assez peu à l'heure actuelle, en dehors des études de Delahaie (2011), Paillard & Vu (2012), Garnier & Savage (2011), Lundquist (2013).

Comme nous ne procédons pas à une enquête des pratiques d'enseignement des enseignants, nous allons effectuer une petite analyse de l'intégration des MD en général, et des séquences lexicalisées à fonction discursive en particulier, dans les méthodes et dans les livres de grammaire de FLE. L'analyse du matériel de travail, en particulier les manuels et les livres de grammaire, est un des moyens susceptibles de nous aider à comprendre le contexte d'enseignement et d'apprentissage de ces éléments lexicaux.

2.2.2.4. Les MD dans quelques manuels de FLE

Face au nombre impressionnant de manuels de FLE qui existent actuellement, le choix n'est pas évident. Nous avons décidé de sélectionner des méthodes assez récentes et utilisées comme supports de travail dans des filières francophones universitaires au Vietnam et dans le Département de français de l'Université nationale de Hanoï (Tableau 2.3). Ce choix est lié au fait que nous allons effectuer dans le chapitre 6 une analyse du corpus de mémoires d'apprenants vietnamophones. Nous nous intéressons particulièrement aux manuels de FLE au niveau intermédiaire et avancé (niveau B2, C1), ce qui correspond au niveau du public visé qui doit atteindre le niveau B2¹¹ du CECRL pour le DELF afin d'être admis en Master dans plusieurs universités françaises. En fait, pour poursuivre les études à l'étranger, les étudiants non natifs se trouvent obligés de passer les certifications en FLE du type DELF/DALF ou TCF pour tester leur niveau en langue étrangère ; ces épreuves de langue respectent les niveaux de référence recommandés par le Cadre Européen de Référence pour les Langues (CECRL). C'est pourquoi les méthodes de français récentes précisent désormais le niveau acquis en fin de formation. En ce qui concerne les MD, d'après le CECRL (2001), il est demandé aux apprenants au niveau B2 d'être aptes à « utiliser un certain nombre limité

¹¹ D'après le CECRL, le niveau B2 décrit un utilisateur indépendant qui « Peut comprendre le contenu essentiel de sujets concrets ou abstraits dans un texte complexe, y compris une discussion technique dans sa spécialité. Peut communiquer avec un degré de spontanéité et d'aisance tel qu'une conversation avec un locuteur natif ne comportant de tension ni pour l'un ni pour l'autre. Peut s'exprimer de façon claire et détaillée sur une grande gamme de sujets, émettre un avis sur un sujet d'actualité et exposer les avantages et les inconvénients de différentes possibilités. » (CECRL, 2000 : 25)

d'articulateurs pour lier ses phrases en un discours clair et cohérent bien qu'il puisse y avoir quelques "sauts" dans une longue conversation » (CECRL, 2001 : 28).

Nous avons sélectionné cinq manuels de niveau B2 (*Le nouvel édito*, *écho B2*, *Tout va bien 3*, *Studio 3*, *Alter ego 4*) et un de niveau C1-C2 (*Alter ego 5*) (Tableau 2.3). Il est à préciser que ces méthodes sont destinées à un large public. Par cette enquête, nous cherchons à répondre aux questions suivantes : 1/ Quelle est l'étiquette qu'attribuent les concepteurs des manuels à cette unité lexicale ? 2/ Par quelle approche les MD sont-ils introduits dans les manuels de FLE ? 3/ Ces méthodes de FLE sont-elles adaptables à notre public FLE ?

Méthodes FLE	Niveau	Édition	Année de parution
Le Nouvel édito	B2	Didier	2012
écho	B2	CLE International	2010
Tout va bien 3	B2	CLE International	2005
Studio 3	B2	Didier	2008
Alter ego 4	B2	Hachette	2007
Alter ego 5	C1-C2	Hachette	2010

Tableau 2.3 : Liste des méthodes de FLE consultées

Parmi les cinq méthodes de FLE de niveau B2 que nous avons consultées, les MD sont souvent appelés connecteurs ou regroupés sous des termes assez larges par exemple *la cause*, *l'opposition*, *la conséquence*, etc. ou bien *l'expression de la cause*, etc. Pour chaque groupe, les apprenants disposent de moyens lexicaux pour exprimer la cause, l'opposition, etc. Par exemple, les apprenants ont sous forme de récapitulatif les moyens pour exprimer la cause avec une conjonction (*comme*, *puisque*, etc.), ou une préposition (*avec*, *sans*, *vu*, etc.). À notre avis, le terme « connecteurs logiques » est un peu étroit et n'évoque pas toutes les fonctions que jouent les marqueurs dans le discours : les fonctions d'introduction des informations, de reformulation, de topicalisation par exemple ; mais nous n'allons pas entrer dans la discussion sur le choix de la terminologie. Nous sommes d'accord sur le fait qu'il faudrait une étape de didactisation de la terminologie, car il n'est pas souhaitable d'introduire des termes linguistiques aux apprenants. En ce qui concerne la façon dont ces « connecteurs » sont introduits dans ces manuels, les apprenants disposent d'exemples et de quelques exercices

d'entraînement, mais il s'agit pour tous de phrases hors contexte. Tous ces outils linguistiques sont présentés sous forme de listes et accompagnés éventuellement d'une brève explication d'emploi. Les marqueurs de reformulation, d'exemplification, d'addition sont pour la plupart absents des manuels de FLE, sauf dans *Le Nouvel édito* B2 où nous avons repéré un tableau récapitulatif des mots de liaison. Il s'agit d'une liste d'expressions qui servent à introduire une séquence, ajouter une information, préciser ou conclure. Ces expressions sont illustrées par des exemples de la langue générale venant de différents genres (revues, romans, émissions de radio, micro-trottoir) et introduites dans des exercices lacunaires ou des exercices de reformulation.

À notre connaissance, il existe peu de méthodes de FLE pour le niveau C1-C2. *Alter ego 5* est destiné au public souhaitant perfectionner son français par différentes activités fondées sur des séquences écrites et orales, très variées en termes de discours, de registres, etc. Dans le livre de l'élève, la partie grammaire est sommaire et représentée sous forme d'un récapitulatif des moyens lexicaux (les auteurs considèrent que les connaissances grammaticales sont acquises à ce niveau) pour laisser la place à d'autres activités de compréhension écrite et orale, production écrite et orale, des exercices d'enrichissement lexical, des techniques pour s'entraîner à des tâches universitaires et administratives (faire un exposé, une synthèse par exemple). Le livre de perfectionnement propose aux apprenants des exercices d'entraînement. Il peut s'agir d'exercices lacunaires, d'exercices de reformulation. Il nous semble dommage qu'il manque des exemples contextualisés et des exercices qui portent sur les différentes valeurs sémantiques des MD, parce qu'il est important, à notre avis, d'apporter aux étudiants de ce niveau davantage d'occasions d'accéder aux subtilités de la langue.

Comme nous l'avons précisé, si cette recherche en thèse porte essentiellement sur les séquences polylexicales dans les écrits scientifiques, nous ne négligeons pas l'enseignement des unités monolexicales, mais ces dernières sont davantage abordées dans les méthodes de FLE et dans les livres de grammaire de FLE que les éléments polylexicaux. En effet, dans les méthodes de FLE pour les étudiants de niveau avancé, *Studio 3*, *Tout va bien 3*, etc., ces unités polylexicales ne font pas encore l'objet d'un enseignement systématique. Elles sont traitées d'une manière succincte dans les activités de lecture ou dans les tableaux récapitulatifs. Par exemple, dans *Studio 3*, pour exprimer la finalité, l'apprenant ne peut trouver qu'une liste de plusieurs mots et locutions sur le but, l'objectif. Selon nous, les étudiants tireront davantage profit de l'apprentissage s'ils sont exposés au contexte d'utilisation des MD, ce qui les aidera à mieux mémoriser l'emploi de chaque expression.

La définition de l'emploi de certains « connecteurs » est parfois trop rapide à notre avis ou elle est absente. Prenons un exemple de la présentation des ressources grammaticales dans *écho B2*, plus précisément pour enseigner les expressions de la conséquence (Livre de l'élève, page 59). Les apprenants disposent d'une liste de différentes expressions pour introduire une conséquence. Les connecteurs suivants n'ont pas d'explication d'emploi : *donc, par conséquent, de ce fait*, d'autres ont une petite définition : *voilà pourquoi, c'est pourquoi, c'est pour cela que, d'où* (introduit une idée déjà formulée), *du coup* (introduit plutôt une conséquence rapide), etc. Il nous semble qu'une telle définition ne suffit pas pour aider les apprenants à faire la différence d'emploi entre les connecteurs. Il faudrait que l'enseignant leur explique mieux ce qui distingue ces connecteurs les uns des autres. Par ailleurs, les méthodes pourraient, selon nous, prendre davantage en compte les avancées de la linguistique et utiliser des exemples contextualisés sans trop faire de théorie. Par exemple, l'utilisation de concordanciers pour l'apprentissage du français pourrait être très bénéfique aux apprenants (Cf. 6.3.2 pour des exemples d'utilisation des concordanciers dans l'enseignement).

Concernant les approches adoptées dans ces méthodes, nous constatons que l'approche sémasiologique reste l'approche la plus fréquente. En fait, l'enseignement des MD ou connecteurs logiques (termes souvent utilisés dans les méthodes de FLE) se résume souvent en la mémorisation des listes de mots et expressions. L'approche onomasiologique est également employée dans la mesure où chaque leçon envisage d'introduire des notions relativement larges comme les expressions de la cause, de la concession, etc.

Ces différents documents ainsi que les exercices dans les manuels de FLE peuvent correspondre à un large public d'apprenants, mais ils ne sont pas très adaptés, nous semble-t-il, aux attentes de nos publics étudiants qui sont amenés à devoir beaucoup rédiger et ont besoin de textes dans un genre spécifique, l'écrit scientifique. Ce public scientifique a des besoins linguistiques spécifiques, pour suivre des cours et rédiger des écrits universitaires en français. Il aura à travailler sur le vocabulaire, les expressions spécifiques dans les écrits scientifiques afin de se familiariser avec cette langue et de pouvoir les réinvestir dans ses productions écrites. Si nous n'excluons pas les exercices lacunaires dans les activités de renforcement de connaissances, il serait plus judicieux à notre avis de laisser réfléchir les apprenants sur la manière dont on peut remplir les « trous ». Dès l'instant où il y a une réflexion, il y a selon nous une pertinence pédagogique.

En définitive, les manuels de FLE fonctionnent bien pour connaître les emplois les plus fréquents des MD, mais il faut également aborder les valeurs sémantiques de chaque

MD. C'est la relation entre les éléments reliés par les MD qui sert à prédire le sens des MD et à distinguer un MD d'un autre. C'est ce que nous allons montrer dans le Chapitre 5 et Chapitre 6. Une telle activité de sensibilisation au plan grammatical est tout à fait possible si l'enseignant fait travailler les apprenants sur un contexte précis et des productions authentiques. De plus, il est indispensable de combiner l'approche sémasiologique et l'approche onomasiologique pour proposer une entrée méthodologique qui tienne compte de tous les aspects linguistiques de l'élément à enseigner. Pour un public universitaire, il est important, nous semble-t-il, de donner l'occasion de travailler dans des contextes précis qui permettront d'éviter de longues explications. Il est utile de donner des moyens lexicaux, mais il serait mieux de montrer aux apprenants dans quel contexte ils peuvent les utiliser.

2.2.2.5. Les MD dans les grammaires de FLE

Nous allons maintenant analyser comment les MD sont traités dans quelques grammaires de FLE très utilisées pour les apprenants (Tableau 2.4).

Grammaire	Editeur	Editeur scientifique	Année de parution
La grammaire des premiers temps (niveau 2)	PUG	Dominique Abry Marie-Laure Chalaron	1999
Grammaire du français	Hachette	Yves Delatour	1991
Grammaire utile du français	Didier	Evelyne Berard Christian Lavenne	1991
Précis de conjugaison	CLE Internationnal	Isabelle Chollet Jean-Michel Robert	2005

Tableau 2.4 : Livres de grammaire de FLE consultés

Notre analyse de ces ouvrages révèle que les MD sont traités d'une manière très disparate. Ils sont groupés sous des appellations générales, par exemple « l'expression de la condition, de la cause, de la finalité » dans : *Grammaire du français*, *Grammaire utile du français* ; ou « les articulateurs » dans *Précis de conjugaison*. Dans les ouvrages de grammaire, les apprenants ont des listes de mots et des expressions correspondant à un concept. Il nous paraît dommage que ces outils linguistiques ne soient pas mieux détaillés. Par

exemple, dans la *Grammaire des premiers temps*, chaque MD est décrit brièvement en une ou deux phrases, les phrases – énoncés exemplaires – tiennent lieu d’explication, mais il n’existe que des exemples hors contexte. Nous préconisons d’utiliser les exemples pour montrer le fonctionnement des connecteurs dans le contexte pour mieux faire ressortir leur sens et leur emploi.

Dans ces grammaires, l’approche sémasiologique est la plus fréquente. Chaque connecteur est accompagné d’une brève définition. Cela peut être utile pour les apprenants qui veulent connaître l’emploi général d’un connecteur, mais il n’est pas suffisant dans tous les cas à notre avis pour aider les apprenants à faire la distinction entre les connecteurs. Les grammaires FLE pourraient regrouper les connecteurs qui partagent la même valeur sémantique et proposer aux apprenants d’y accéder par valeurs sémantiques, puis leur donner des moyens lexicaux pour exprimer ces valeurs.

En définitive, ces grammaires peuvent bien convenir à un public non spécialiste et non universitaire. Pourtant, il y manque des utilisations contextualisées pour nos publics universitaires qui ont besoin de beaucoup plus d’entraînement pour une bonne utilisation des éléments phraséologiques spécifiques. De plus, il nous semble que le sens de certains emplois des MD dans ces grammaires n’est parfois pas clair. Il manque l’emploi discursif et les valeurs sémantiques des MD. Certes, un public non universitaire n’a pas forcément besoin d’une telle précision au niveau sémantique, mais il est important selon nous d’aborder différentes valeurs sémantiques de ces MD pour un public universitaire, ce qui explique la nécessité de nous orienter vers de nouvelles pistes pour mieux adapter l’enseignement/apprentissage de ces unités lexicales à un public universitaire, comme le *Français scientifique et technique* de Eurin Balmet & Henao de Legge (1992) et du *Français sur Objectif Universitaire* de Mangiante & Parpette (2012).

En bref, nous avons relevé dans cette partie les erreurs principales dans l’utilisation des organisateurs textuels dans les productions écrites des apprenants, sur les plans syntaxique, sémantique et discursif. Nous avons également effectué un petit diagnostic dans les manuels de FLE et les livres de grammaire FLE pour voir comment les MD y sont introduits. Le public universitaire a besoin d’un lexique spécifique pour rédiger dans un genre spécifique avec des règles conventionnelles. Il nous semble indispensable, dans un premier temps, de sensibiliser les apprenants aux différents rôles de ces MD pour le guidage du lecteur, ou pour la structuration du texte par des exemples précis. Cela éviterait de recourir à des explications linguistiques et/ou grammaticales parfois complexes. Dans un deuxième temps, il est important de traiter à la fois le fonctionnement syntaxique et sémantique des MD.

Par l'analyse linguistique, nous allons montrer que les deux éléments liés par le MD sont cohérents et que le MD est le point de rencontre des valeurs sémantiques.

2.2.3. Les marqueurs discursifs comme objets d'enseignement et d'apprentissage

À notre connaissance, peu de recherches existantes traitent de la didactisation des MD (Delahaie, 2011 ; Paillard & Vu, 2012). Delahaie (2011) considère les MD comme un objet d'enseignement particulièrement pertinent pour le public Erasmus. Elle travaille sur des MD de l'oral, mais qui relèvent à la fois du français académique et du français quotidien. Selon l'auteure, ce public Erasmus présente une forte hétérogénéité en ce qui concerne le profil linguistique et les objectifs d'apprentissage. Tout cela explique la nécessité de travailler sur des compétences transversales et de s'intéresser aux objectifs et aux objets d'enseignement transversaux parmi lesquels les MD constituent un outil linguistique à introduire dans les cours destinés aux étudiants Erasmus. Au lieu de fournir simplement aux étudiants une liste de connecteurs logiques, Delahaie propose de traiter des MD par paire afin de mettre en lumière les propriétés syntaxiques et sémantiques de chaque unité lexicale, ce qui nous paraît tout à fait pertinent. Par exemple, pour faire la distinction entre *en fait* et *de fait*, l'auteure ne cherche pas à donner des explications compliquées, mais procède plutôt à une explication sémantique illustrée par des exemples contrastifs. Cette démarche nous semble intéressante en ce qu'elle permet de mettre l'accent sur les MD les plus fréquents dans le discours et elle est adaptable aux séquences d'enseignement à destination d'autre public, y compris le public scientifique.

À l'instar de Delahaie (2011), nous n'envisageons pas de proposer aux étudiants la liste des connecteurs logiques ou un cours sur l'utilisation des termes qui expriment la cause, la conséquence, la concession, ce qui n'est guère intéressant sur le plan didactique. En outre, des explications compliquées et spécifiques des fonctions de chaque marqueur n'intéressent pas les étudiants ; en revanche, des exemples authentiques et précis dans les écrits scientifiques peuvent tout à fait leur fournir un terrain d'entraînement, l'objectif étant que les étudiants puissent réinvestir ultérieurement toutes ces connaissances dans leurs écrits, il serait donc pertinent de leur donner à voir des exemples d'utilisation.

Le projet *Inventaire raisonné des marqueurs discursifs du français*¹² (Paillard & Vu, 2012) attire également notre attention par ses objectifs linguistique et didactique. Il envisage de proposer des descriptions linguistiques des MD de la langue française afin de répondre aux problèmes en didactique du FLE. Sur le plan linguistique, les chercheurs proposent une description systématique des MD. L'objectif est de repérer les propriétés et les régularités syntaxiques et sémantiques qui sont susceptibles d'être réinvesties dans l'enseignement du FLE. Vingt-et-un MD ont été analysés et certains marqueurs sémantiquement proches sont regroupés : 1. *en fait, au fait, de fait* ; 2. *en revanche, au contraire, par contre* ; 3. *en effet, effectivement* ; 4. *en réalité, réellement* ; 5. *enfin, finalement* ; 6. *d'ailleurs, par ailleurs* ; 7. *pourtant, cependant* ; 8. *quand même, tout de même*, 9. *tiens*, 10. *déjà* ; 11. *alors*. La mise en confrontation des MD de proximité sémantique nous semble très pertinente, car les marqueurs de même catégorie comme *en fait, de fait, au fait* sont souvent considérés comme des synonymes.

Par des analyses linguistiques assez fines, les auteurs proposent différentes étapes pour la transposition didactique des MD du français, ainsi que quelques exemples de fiches pédagogiques pour expliquer la manière de les introduire dans l'enseignement du FLE. Bien que les MD appartiennent à des catégories grammaticales très diversifiées, Paillard & Vu (2012) ont relevé les régularités sémantiques et syntaxiques propres à ces mots, ce qui facilitera par la suite une étape de didactisation. Nous allons décrire en détail la typologie de Paillard & Vu (*ibid.*) ultérieurement.

En fin de compte, la modélisation des MD pour une perspective didactique doit tenir compte de toutes ces caractéristiques syntaxiques et sémantiques ainsi que de leurs propriétés distributionnelles et de leur portée qui peut être variable. Comme nous l'avons précisé précédemment, l'utilisation didactique des MD ne pourrait se faire sans les connaissances syntaxiques et sémantiques de ces unités lexicales. Nous analysons quelques typologies des MD afin de parfaire par la suite notre classement de ces MD pour leur enseignement en FLE.

¹² Ce projet, piloté par Mme Vu Thi Ngan, ESLE, Université Nationale de Hanoï, et M. Denis Paillard, Laboratoire de Linguistique Formelle, Université Paris Diderot (2009-2011), a été soutenu matériellement par l'Agence Universitaire de la Francophonie pour la région Asie-Pacifique. Ce projet résulte de la coopération de l'Université Nationale de Hanoï, l'Université de Hanoï, l'Université Paris Diderot et le Laboratoire de linguistique formelle de l'Université Paris Diderot.

2.3. Marqueurs discursifs : typologie

Plusieurs typologies des MD ont vu le jour, mais aucune d'entre elles ne fait réellement autorité. Les typologies présentées ci-dessous nous intéressent, soit par leur approche descriptive des MD qui est susceptible d'être réadaptée à l'enseignement aux apprenants FLE, soit par leur orientation didactique. Elles proposent des appellations très diverses des MD, mais toutes les descriptions sont d'ordre pragmatico-fonctionnel.

2.3.1. Typologie de Riegel

Dans la *Grammaire méthodique du français*, Riegel *et al.* (2009) définissent les connecteurs comme « des termes de liaison et de structuration ». (*ibid.* : 1044). Les connecteurs englobent des conjonctions de coordination (*car, donc, et, etc.*), des adverbes (*pourtant, ensuite, autrement dit, en tout cas*), des groupes prépositionnels (*en ce qui concerne, quant à*). Ils contribuent à la « structuration du texte en marquant des relations entre les propositions ou entre les séquences qui composent le texte et en indiquent les articulations du discours » (*ibid.* : 1044). Cette définition résume bien les principaux rôles de ces connecteurs. Riegel *et al.* (2009) relèvent différentes fonctions des connecteurs, dont le rôle d'organiseurs textuels, la fonction énonciative et la fonction pragmatique.

En se basant sur les fonctions des connecteurs, ils distinguent les organisateurs textuels des *marqueurs de prise en charge énonciative* et des *connecteurs argumentatifs*. Les organisateurs textuels comprennent aussi bien les connecteurs temporels et spatiaux que les organisateurs de texte et partagent le même objectif de structurer le texte. Les connecteurs temporels (*alors, après, ensuite, puis, etc.*) et les connecteurs spatiaux (*en haut, en bas, au-dessus, en dessous, etc.*) s'emploient pour marquer une succession chronologique ou une localisation spatiale. Comme Riegel *et al.* (*op. cit.*) l'ont évoqué, ce type de connecteur ordonne la réalité référentielle. À l'instar de Tutin (2014), nous intégrons plutôt les connecteurs temporels et spatiaux à la catégorie des marqueurs à fonction référentielle qui visent à décrire les phénomènes, les objets, les réalités de la vie. De l'autre côté, les organisateurs de la mise en texte structurent la progression du texte et englobent les connecteurs énumératifs (*également, en plus, de plus, en outre, etc.*), les marqueurs de topicalisation (*quant à, à propos de etc.*) et les marqueurs d'exemplification et d'illustration (*en particulier, entre autres, etc.*).

Selon Riegel *et al.* (*op. cit.*), les marqueurs de prise en charge énonciative contribuent au fonctionnement du discours. Tandis que ceux-ci insistent sur l'apparition d'un point de vue du locuteur, les organisateurs de la mise en texte participent au découpage du texte en

différentes parties. Ils englobent les marqueurs de point de vue (*d'après, selon, pour*), les connecteurs de reformulation (*c'est-à-dire, autrement dit, en d'autres termes, bref*) et les marqueurs de structuration de la conversation. Ces derniers jouent un rôle important dans la structuration des discours oraux, y compris les marqueurs de structuration proprement dits (*bon, ben, alors, etc.*) et les phatiques (*tu sais, tu vois, euh, etc.*). Comme nous nous intéressons dans cette recherche uniquement au discours écrit, nous excluons tous les connecteurs appartenant au discours oral ou aux séquences dialogales comme les marqueurs de structuration et de la conversation.

La dernière catégorie comprend les connecteurs argumentatifs qui servent à marquer les relations entre les énoncés, les séquences, ainsi qu'une argumentation. Riegelet *al. (ibid.)* ont relevé des paires de connecteurs comme opposition – concession (*pourtant, quand même, au contraire, certes...mais, etc.*), explication et justification (*car, puisque, en effet, etc.*), la complémentation (*or, d'ailleurs, non seulement... mais encore*), et la conclusion (*donc, en tout cas, somme toute, etc.*).

Pour finir, Riegelet *al.* ont présenté une synthèse assez claire et détaillée des connecteurs en français dans le discours écrit et oral. Cette classification nous permet d'avoir une vue globale des connecteurs en langue générale et nous a inspirée par la suite dans la conception de notre typologie.

2.3.2. Typologie de Charolles

Dans son étude sur l'encadrement du discours, Charolles (1997) a pour objet les marques de nature linguistique, marques de cohésion « qui codent des instructions relationnelles plus ou moins spécifiques » (*ibid.* : 2). Les marqueurs de cohésion constituent un ensemble linguistique hétérogène. Charolles en distingue deux grands types : 1. les relations de connexion entre deux unités adjacentes ou proches, classe qui comprend deux sous-classes dont les anaphores et les connecteurs ; 2. les relations d'indexation indiquent que plusieurs propositions du texte entretiennent un même rapport avec un certain critère. Elles sont induites par des adverbiaux cadratifs ou bien les expressions introductrices de cadre dont la portée est plus ou moins large (limitée à une phrase ou plusieurs propositions) et regroupées sous la notion de *cadres*.

Charolles (2003) distingue quatre types de cadres de discours, qui portent soit sur l'énoncé soit sur l'énonciation à savoir : 1. cadres thématiques ; 2. cadres organisationnels ; 3. univers de discours ; 4. cadres qualitatifs. Les univers de discours et les cadres qualitatifs sont

les deux types de cadratifs qui caractérisent l'énoncé. D'un côté, les *cadres qualitatifs* (ou *cadratifs véridictifs*) sont introduits par des expressions qui apportent des précisions sur les aspects qualitatifs des états de chose (*pour le plus grand bonheur*). D'un autre côté, les *univers de discours* (ou bien *cadratifs non véridictifs*) sont des circonstances spécifiées sous forme d'adverbe (*en France, chaque année*). Dans notre recherche, nous ne retenons pas ces deux types de cadres, car d'après la typologie des séquences polylexicales proposée par Tutin (2014), les cadres qualitatifs renvoient plutôt aux séquences à fonction interpersonnelle et modale, tandis que les univers de discours font partie des séquences à fonction référentielle.

En outre, deux autres types cadratifs servent à caractériser l'énonciation dont les cadres thématiques et les cadres organisateurs. Les *cadres thématiques* servent à spécifier un champ de connaissance, un domaine de représentation. Ils mettent l'accent sur la dimension illocutoire de l'énonciation, comme *concernant, quant à*, etc. Enfin, les *cadres organisateurs* sont des cadres introduits par des expressions qui portent sur des aspects métalinguistiques de l'énonciation et de l'organisation du discours (*bref, en résumé*, etc.) (Charolles, *ibid.*) Ces types cadratifs correspondent aux séquences lexicalisées à fonction discursive précédemment décrites. La Figure 2-1 sert à situer ces deux cadratifs par rapport aux autres types de cadres du discours.

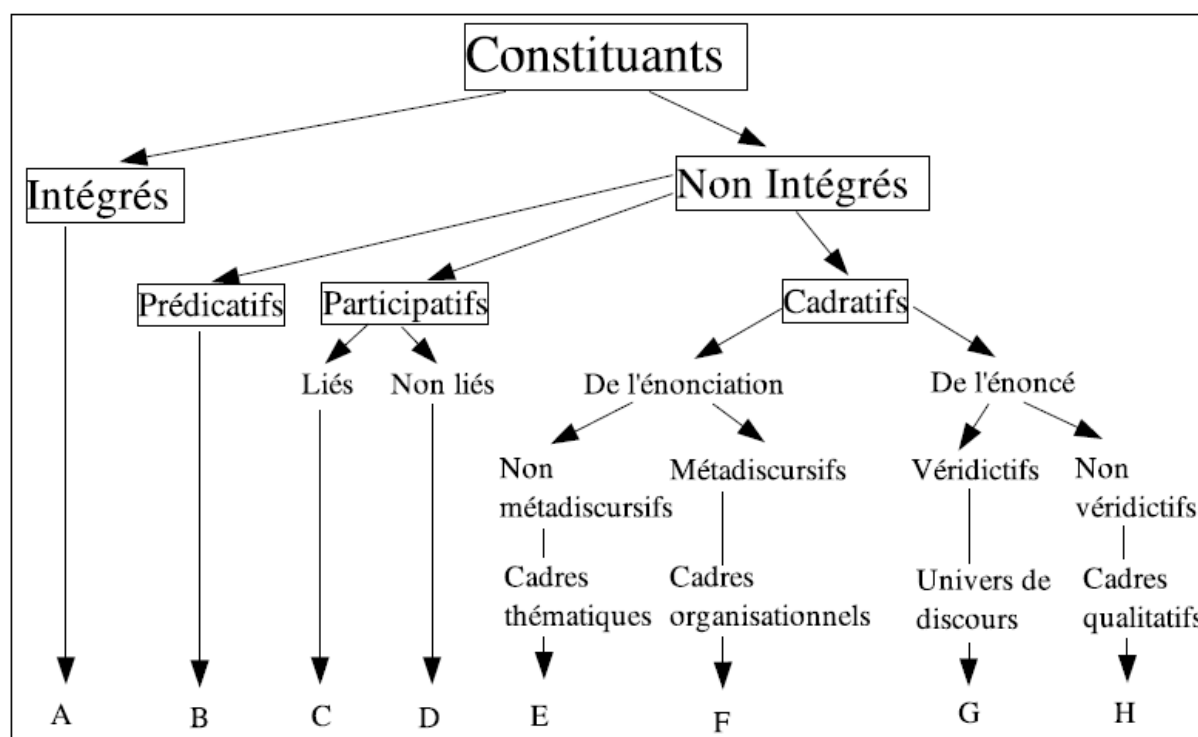


Figure 2-1 : Place des cadratifs parmi l'ensemble des constituants (Charolles, 2003)

Ces cadres du discours nous semblent très intéressants en ce qu'ils permettent de relever des segments discursifs localement cohérents. Le repérage des cadres du discours se fait par les introducteurs de cadres qui sont en général des circonstants. Sur le plan didactique, une activité de repérage des cadres du discours grâce aux introducteurs de cadres aidera les apprenants à relever les différentes structures discursives et à développer des compétences discursives.

2.3.3. Typologie de Siepmann

Siepmann (2005,2007) a relevé qu'il n'y avait pas assez de recherche en lexicographie portant sur les unités polylexicales à valeur discursive malgré leur importance pour le traducteur. Dans son approche didactique, il propose de classer les principaux types de MD polylexicaux selon des critères fonctionnels et d'introduire des listes d'exemples des marqueurs susceptibles de servir à leur enseignement. Les marqueurs polylexicaux sélectionnés se caractérisent par leurs propriétés d'ordre pragmalinguistique et lexicologique. Sur le plan pragmalinguistique, les MD consistent en des expressions de la langue naturelle dont la fonction de base est de faciliter le processus d'interprétation des relations de cohérence entre une unité particulière du discours et d'autres qui entourent les aspects de la situation communicative (Risselada et Spooren, 1998, cité par Siepmann, 2007).

Du point de vue lexicologique, les unités phraséologiques présentent un certain degré de figement. Cela dit, les MD polylexicaux se distribuent sur les quatre premiers niveaux d'un continuum à cinq niveaux allant du plus figé au moins figé et ils n'appartiennent pas aux combinaisons libres. Siepmann distingue les marqueurs polylexicaux en deux groupes en fonction de leur fréquence, « les marqueurs de premier niveau » (*first-level discourse markers*), c'est-à-dire les marqueurs les plus fréquents (*en effet, en revanche, etc.*), et les « marqueurs de deuxième niveau » (*second-level discourse markers*) les moins fréquents (*Force est de constater que, pour commencer, etc.*). Il prête un intérêt particulier aux marqueurs du deuxième groupe qui se composent des locutions et des collocations de fréquence moyenne et permettent de « faciliter le processus d'interprétation des relations de cohérence entre différents segments textuels et/ou des aspects de la situation d'énonciation » (Siepmann, 2007 : 125).

Siepmann (*ibid.*) distingue 22 groupes de marqueurs polylexicaux à valeur discursive. Chaque catégorie fonctionnelle s'accompagne d'une brève définition et s'illustre par quelques

exemples. Cette liste de marqueurs polylexicaux nous semble assez longue et complexe, et certaines catégories pourraient être regroupées. À titre d'exemple, les marqueurs d'annonce, d'introduction thématique et de définition sont susceptibles d'être groupés dans une catégorie de marqueurs « qui apporte une nouvelle information » (*ibid.* : 129). Ou bien les marqueurs d'énumération, les résumés et les charnières de terminaison et les marqueurs d'introduction thématique se distinguent uniquement par la position du segment textuel où ils interviennent dans le texte. Si les marqueurs d'introduction thématique introduisent un nouveau thème et se trouvent au début du texte pour faciliter l'entrée en matière, les résumés et les charnières de terminaison introduisent de leur côté le dernier terme d'une série et marquent la fin. Ces types de marqueurs pourraient être regroupés dans la catégorie des marqueurs énumératifs par exemple, car ils renvoient tous à la notion de l'énumération.

Il est également à noter que la définition de Siepmann est plus large que la nôtre dans la mesure où nous n'intégrons pas les marqueurs liés à la situation d'énonciation dans la classe des interpersonnels (et non discursifs) (*il est communément admis que nous renvoyons sur ce point à, pour une plus ample discussion*). L'approche didactique de Siepmann nous paraît intéressante, en particulier parce qu'il prête un intérêt particulier aux marqueurs de deuxième niveau, c'est-à-dire aux locutions de fréquence moyenne.

2.3.4. Typologie de Paillard et Vu

Les objectifs de Paillard & Vu (2012) sont à la fois d'ordre linguistique et didactique. Les descriptions des propriétés syntaxiques, sémantiques et distributionnelles des MD seront réinvesties dans l'enseignement au public FLE. D'après Paillard & Vu (2012), le terme « discursif » ne signifie pas que ces marqueurs interviennent dans la construction du discours. Par « discursif », ils entendent « le fait que ces marqueurs confèrent un statut discursif (énonciatif) particulier à la séquence p correspondant à leur portée pour ce qui est d'exprimer un état de choses Z » (*op.cit.* : 19). Cette définition permet de différencier les MD des connecteurs qui ont pour fonction première de mettre en relation deux entités (deux mots, deux propositions).

Paillard et Vu distinguent quatre classes de MD du point de vue de la sémantique discursive¹³ : 1/ les mots du discours ; 2/ les particules énonciatives ; 3/ les modalisateurs et 4/ les mots du dire. Premièrement, **les mots du discours** sont des éléments utilisés, soit pour

¹³ D'après Paillard & Vu (2012), la notion « sémantique discursive » recouvre les différents facteurs qui interviennent dans la constitution d'un dire. Cette notion n'est pas spécifique aux MD ; elle renvoie à « l'activité de langage dans sa généralité » (*ibid.* : 14).

qualifier l'événement exprimé par la séquence, soit pour donner le point de vue sur l'état des choses. Les auteurs distinguent deux sous-classes dont les MD catégorisants (*malheureusement, naturellement, etc.*) et les MD de point de vue (*en réalité, vraiment, réellement, effectivement, etc.*). Deuxièmement, **les particules énonciatives** apportent de nombreuses informations à l'énoncé dans le contexte gauche. On distingue les particules de type 1, formées à l'aide de prépositions et d'adverbes (*pourtant, cependant, par contre*), et les particules de type 2, appartenant au discours oral (*bien, bon, déjà, etc.*). Troisièmement, **les modalisateurs** présentent une forme dont le sens entretient des rapports étroits avec la notion d'indétermination (*en quelque sorte, genre, quoi*). Enfin, **les mots du dire** correspondent à une forme de désengagement du sujet et actualisent une distance entre un « vouloir dire » et l'énoncé (*disons, pour ainsi dire, soit disant, à vrai dire, etc.*).

La typologie de Paillard & Vu (*ibid.*) est constituée à la fois des MD du discours écrit et oral. Dans notre recherche, nous retenons uniquement les particules énonciatives et une partie des mots du dire qui se trouvent dans le discours écrit. Les termes appartenant au discours oral sont exclus. Les mots du discours et les modalisateurs font partie selon nous des séquences polylexicales à fonction modale qui expriment une attitude du scripteur. Les catégories proposées par Paillard & Vu (*ibid.*) nous semblent un peu larges. Les auteurs ne proposent pas de sous-catégories qui seraient susceptibles de mieux regrouper les MD. Comme nous l'avons préconisé antérieurement, leur travaux font partie des rares recherches, à notre connaissance, qui allient les analyses linguistiques et les descriptions didactiques. Pourtant, les fiches pédagogiques que les auteurs proposent suivent uniquement une approche sémasiologique. À notre avis, en regroupant les MD par leurs fonctions, dans une approche onomasiologique, on pourrait mieux guider les apprenants dans leur apprentissage des MD, ce qui les aiderait à mieux mettre en comparaison des MD sémantiquement proches par exemple.

Conclusion

Pour résumer, dans cette partie, nous avons décrit notre objet de recherche, les MD polylexicaux ou les séquences lexicalisées à fonction discursive. Du fait de leur structure syntaxique, ces unités lexicales posent des difficultés syntaxiques et sémantiques aux étudiants non natifs qui doivent rédiger leur mémoire en français. L'analyse que nous avons effectuée sur des manuels et des livres de grammaire de FLE montre que les manuels à visée généraliste ne peuvent pas encore répondre à nos attentes. Il nous paraît nécessaire de prendre

en compte les aspects syntaxiques et sémantiques des MD, ainsi que des exemples contextualisés pour donner plus d'occasions d'exposer les apprenants à un lexique spécifique dans les écrits scientifiques. Enfin, l'analyse de différentes typologies des MD nous confirme l'importance d'une classification des séquences lexicalisées à fonction discursive. Pour notre typologie, nous allons nous baser sur les fonctions que jouent ces éléments linguistiques dans le discours. Une telle classification apparaît compatible avec une approche onomasiologique que nous souhaitons mettre en place par la suite au plan didactique.

Chapitre 3.

Méthodologie

Comme nous l'avons précisé, notre objectif de travail est à la fois d'ordre linguistique et didactique. Sur le plan linguistique, nous souhaitons caractériser un lexique transversal, des séquences lexicalisées à fonction discursive grâce à une étude sur un corpus d'écrits scientifiques. Sur le plan didactique, nous souhaitons tirer profit des avantages du corpus pour développer des réflexions didactiques autour de ces éléments phraséologiques, essentiellement pour l'aide à la rédaction scientifique pour des locuteurs non natifs (désormais LNN).

Nous décrivons dans ce chapitre la construction d'une méthodologie adaptée à nos objectifs de recherche, à commencer par la constitution du corpus. Nous allons expliciter par la suite les critères de sélection des séquences lexicalisées à fonction discursive. Enfin, nous présenterons les approches que nous adoptons pour l'analyse linguistique et pour la partie didactique.

3.1. Une approche inscrite dans la linguistique de corpus

Notre travail de recherche se situe dans le cadre de la linguistique de corpus, dont les origines se trouvent dans la tradition anglo-saxonne, pour des objectifs de l'enseignement de la langue. En s'inscrivant dans une perspective didactique, l'analyse linguistique de nos MD polylexicaux doit être fondée sur des exemples authentiques, c'est-à-dire un corpus composé de documents attestés. Sinclair (1991), chercheur pionnier de la linguistique de corpus moderne, a défini le corpus comme une collection de données qui sert d'échantillon du langage¹⁴ (Sinclair, 2005). Cette définition du corpus révèle certes la propriété d'authenticité du corpus, mais elle nous semble un peu restreinte, en ce sens qu'elle ne précise pas les

¹⁴ en anglais : « a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research » (Sinclair, 2005 : 16)

critères prioritaires dans la construction du corpus. Comme l'ont mentionné Habert *et al.* (1998), le corpus doit résulter d'un « regroupement raisonné » (*ibid.* : 35). Nous rejoignons Habert *et al.* (1997) quand ils soulignent l'importance de construire le corpus en respectant les objectifs théoriques de la recherche, ainsi que « les critères établis en référence aux hypothèses qu'il doit étayer empiriquement » (*ibid.* : 144)

Le corpus ne peut être considéré comme un sac de mots ou une archive de textes, et la constitution du corpus doit répondre à des critères précis. Les propriétés du corpus proposées par Rastier (2002) correspondent bien à nos objectifs d'ordre linguistique et didactique. Selon lui,

un corpus est un regroupement structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages, et rassemblés : (i) de manière théorique réflexive en tenant compte des discours et des genres, et (ii) de manière pratique en vue d'une gamme d'applications. (*ibid.* : 2)

En fait, dans notre travail, le corpus a été créé d'abord à des fins didactiques. Nous avons choisi plusieurs disciplines en SHS pour toucher davantage d'étudiants de différentes disciplines. Ensuite, ce corpus nous servira de ressources textuelles pour des analyses syntaxiques et sémantiques des séquences lexicalisées à fonction discursive les plus représentatives dans les écrits scientifiques.

Nous allons maintenant préciser les différentes étapes de constitution de notre corpus.

3.2. Constitution du corpus

Notre corpus, d'environ 2 680 000 mots, rassemble des articles relevant de 10 disciplines en SHS (linguistique, psychologie, sciences de l'éducation, économie, sciences politiques, anthropologie, histoire, géographie, sciences de l'information et de la communication (SCI), sociologie). Les sous-corpus correspondant à chaque discipline recensent chacun 30 articles, sauf en psychologie et en sciences du langage avec uniquement 25 articles. Ayant pour but de construire une phraséologie transdisciplinaire de base – un lexique transversal aux disciplines scientifiques, notre corpus doit représenter la diversité disciplinaire tout en prenant en compte l'homogénéité des textes retenus. Comme l'ont dit Habert *et al.* (1997), le corpus doit présenter une certaine diversité interne pour limiter les risques que comporte toute entreprise de généralisation. Précisons que nous n'envisageons pas une relative similitude en matière de volume, les articles de certaines disciplines étant bien plus importants en nombre de mots : par exemple, les articles d'histoire sont en moyenne

deux fois plus volumineux que les articles d'anthropologie ou de sciences de l'information (voir Tableau 3.1).

Une description détaillée de la composition de notre corpus se trouve dans l'*Annexe 2 : Liste des revues en SHS*, où figure la liste des articles retenus pour chaque discipline. Les articles sont présentés au moyen de leurs références bibliographiques et suivis de leur numérotation correspondante dans Scientext afin de faciliter toute recherche ou vérification.

Le nombre de mots dans chaque sous-corpus apparaît dans le tableau suivant :

Discipline	Nombre d'articles	Nombre de mots
Linguistique	30	147 255
Psychologie	26	200 683
Sciences de l'éducation	25	195 181
Economie	30	217 390
Sciences politiques	30	424 251
Anthropologie	30	260 068
Histoire	30	457 859
Géographie	30	236 028
Sciences de l'information et de la communication	30	223 916
Sociologie	30	316 180
Total	291	2 678 811

Tableau 3.1 : Liste des disciplines en SHS dans notre corpus

Étant inscrit dans la continuité du projet Scientext¹⁵ « Corpus et outils pour étudier le positionnement et le raisonnement de l'auteur dans les écrits scientifiques », notre travail de

¹⁵ Projet ANR, piloté par Francis Grossmann et Agnès Tutin du laboratoire LIDILEM, Université Stendhal Grenoble 3 (2006-2010). Site de Scientext : <http://scientext.msh-alpes.fr/scientext-site/spip.php?article1>

recherche bénéficie de l'expérience des travaux précédents, en l'occurrence la constitution de quatre sous-corpus en linguistique, psychologie, économie et sciences du langage. La sélection des articles pour les six autres disciplines a été soumise à différents critères afin d'assurer l'homogénéité et la fiabilité de notre corpus concernant le choix du genre d'écrits scientifiques et les sources des textes sélectionnés.

3.1.1. Choix du genre

Parmi les genres d'écrits scientifiques souvent évoqués, à savoir l'article, la thèse, l'Habilitation à Diriger des Recherches, nous nous focalisons tout particulièrement sur les articles de recherche pour nos différents objectifs linguistique, méthodologique et didactique. En effet, ce genre a fait l'objet de nombreuses recherches en anglais (Swales, 1990 ; Teufel, 1999), en français (Berthelot, 2003 ; Pecman, 2004 ; Rinck, 2006) en raison de son statut emblématique dans l'activité scientifique (Bazerman, 1988 cité par Rinck, 2010). Malgré des variations culturelles et disciplinaires, les articles de recherche partagent une structure générale qui vise à répondre à une problématique interne par des argumentations, que ce soit un article où l'auteur présente son travail empirique ou un article argumentatif où l'auteur manifeste son point de vue par rapport aux autres recherches.

En outre, l'article scientifique a la caractéristique d'un genre très structuré (Poudat, 2006 ; Rinck, 2006), il « obéit à des règles et à des codes particuliers au niveau du contenu et de la forme » (Boure, 1998 : 107, cité par Poudat, 2006 : 51). Il peut se diviser en sections et séquences identifiables. Dans le corps de l'article, on retrouve différentes sections comme l'introduction (contexte de recherche, etc.), le développement (organisé ou non selon une structure IMRaD), la conclusion (perspectives, bilan, etc.). Nous nous inspirons du modèle proposé par Teufel (1999) sur les « zones argumentatives » du texte (en anglais : *argumentative zoning*), c'est-à-dire des structures discursives de l'article scientifique pour développer les compétences de compréhension écrite et de production écrite des apprenants. Nous allons montrer que, grâce aux séquences lexicalisées à fonction discursive, les apprenants peuvent rapidement retrouver les structures discursives. De plus, notre corpus d'articles scientifiques nous sert de modèle et apporte des exemples authentiques pour sensibiliser les apprenants aux séquences polylexicales, en l'occurrence les séquences lexicalisées à fonction discursive, comme on le verra au chapitre 6. Comme notre corpus recouvre plusieurs disciplines, l'article de recherche est le plus adapté en raison de sa taille et de sa structure très claire.

3.1.2. Sources des articles

Pour nous aider dans la constitution du corpus d'articles scientifiques, nous nous sommes basée sur un classement des revues en SHS élaboré sur des critères scientifiques. En ce qui concerne les SHS, deux classements des revues scientifiques font autorité : celui de l'AERES¹⁶ *Agence d'évaluation de la recherche et de l'enseignement supérieur*, et le projet ERIH¹⁷ *European Reference Indexe for the Humanities* du CNRS. Nous nous intéressons particulièrement à la classification des revues en SHS de l'AERES pour différentes raisons. Il s'agit d'abord d'un classement effectué en France pour les chercheurs français. Ensuite, il est facilement accessible. En particulier, par manque de connaissances de toutes les revues de toutes les disciplines en SHS, nous avons besoin d'accéder à un classement rigoureux de ces revues. L'AERES met en œuvre des critères bien précis concernant la communauté scientifique, le chercheur, le niveau d'audience. Les listes sont établies par champ disciplinaire par des comités d'experts comprenant des représentants des instances d'évaluation et des documentalistes. La classification A, B, C correspond à l'étendue du rayonnement des revues. Une production scientifique de rang A assure la qualité des articles publiés. C'est pourquoi nous avons fait une présélection des revues classées A par l'AERES¹⁸. Afin d'assurer une certaine homogénéité dans notre corpus, nous nous sommes restreinte à 2 à 3 revues par discipline. En ce qui concerne la diversité du corpus, nous avons choisi des articles de différents numéros et des auteurs différents en vue d'une variation thématique des articles, car chaque numéro d'une revue traite une question spécifique des sous-domaines de chaque discipline.

3.1.3. Périodes de rédaction des articles

Notre corpus contient des articles rédigés entre 1997 et 2006 pour la linguistique, la psychologie, l'économie et les sciences du langage ; et entre 2007 à 2011 pour les autres

¹⁶ Un rapport des évaluations des revues, des établissements, des unités de recherche est rendu public et actualisé chaque année depuis 2008, dont la liste des revues en SHS. Consulté le 22 octobre 2014 : <http://www.aeres-evaluation.fr/index.php/Publications/Methodologie-de-l-evaluation/Listes-de-revues-SHS-sciences-humaines-et-sociales>. L'AERES est également évaluée en 2011 par l'EQAR European Quality Assurance Register for Higher Education, ce qui prouve de sa notoriété à l'échelle mondiale.

¹⁷ Projet ERIH mis en œuvre par l'*European Science Foundation* et le consortium HERA pour classer des revues en A, B, C dans le domaine des sciences humaines. La liste des revues classées se compose de 14 champs disciplinaires et est disponible à l'adresse : <http://www.esf.org/research-areas/humanities/erih-european-reference-index-for-the-humanities.html>

¹⁸ Il s'agit de la liste de revues en Sciences Humaines et Sociales publiées en juillet 2008.

disciplines. Cette différence de période de rédaction est due à la disponibilité des articles lors de la constitution du grand corpus Scientext, car la première phase du projet a démarré en 2007, tandis que notre travail d'enrichissement de Scientext date de début 2011. Tous les articles sont entiers et la majorité des textes sont consultables en ligne. Les articles des six disciplines (histoire, géographie, anthropologie, sciences politiques, sociologie, sciences de l'information et de la communication) ont été extraits du site CAIRN¹⁹, un bouquet de revues en SHS en partie disponibles en ligne gratuitement.

3.1.4. Langue

Étant donné que notre objectif est d'étudier la phraséologie transdisciplinaire en français, nous prenons en compte uniquement des revues publiées en France et les articles devraient être rédigés en langue française. Toutes les traductions sont donc écartées afin de garantir la qualité linguistique de chaque article.

3.1.5. Formats et annotations du corpus

Afin d'obtenir de riches informations pour nos requêtes sur la phraséologie transdisciplinaire de base dans les écrits scientifiques, nous avons recouru à des techniques du TAL, à savoir l'étiquetage morpho-syntaxique et l'analyse syntaxique. Annoter consiste à « ajouter de l'information aux données langagières » (Leech, 2005) (notre traduction, « the practice of adding interpretative linguistic information to a corpus »). En amont du processus d'annotation, le corpus a été étiqueté morpho-syntaxiquement et analysé syntaxiquement par Syntex, qui utilise l'étiquetage de Treetagger²⁰.

Notre corpus a été annoté à trois niveaux : au niveau de la structure, au plan morphologique et au plan syntaxique. En premier lieu, une annotation structurelle avec le standard XML TEI Lite, qui indique les parties textuelles (introduction, conclusion, notes de bas de page, etc.) (Falaise, 2014). En deuxième lieu, l'annotation morphologique et syntaxique a été faite à l'aide de Syntex²¹ – un analyseur syntaxique de surface conçu par Bourigault (2007). L'annotation morphosyntaxique sert à attribuer une étiquette à la forme (c'est-à-dire le mot effectivement trouvé dans le texte), au lemme (c'est-à-dire le mot non fléchi, tel qu'on le trouverait dans une entrée de dictionnaire), à chaque partie du discours (nom, verbe, adjectif, etc.) et aux traits morphosyntaxiques (nombre, genre, temps, mode,

¹⁹ Site du Cairn, consultable en ligne: <http://www.cairn.info/>

²⁰ Treetagger (Helmut Schmid, University of Stuttgart)

²¹ Site de Syntex : <http://w3.erss.univ-tlse2.fr/membres/bourigault/syntex.html>

traits divers (défini/indéfini, relatif, réfléchi, etc.)). Le corpus comporte également des annotations syntaxiques en dépendances qui permettent d'identifier différents phénomènes linguistiques et notamment les relations syntaxiques. Cela nous permet d'extraire les éléments reliés syntaxiquement, comme les collocations.

Constitution du corpus

Afin de faciliter l'exploitation du corpus à l'aide des outils de traitement linguistique, les articles sont récupérés sous format électronique, précisément le format HTML²², étant donné son caractère explicite, et sa quasi-comptabilité avec XML. Ces textes ont ensuite été convertis dans un format unique facilement manipulable, le format XML grâce à un script Perl²³ (Practical Extraction and Report Language)²⁴. Précisons aussi que la conversion des textes de HTML au XML²⁵ complique leur « nettoyage », car suite au reformatage, les textes contiennent un grand nombre de coquilles et une grande partie du processus a été faite manuellement. En vue d'accélérer le nettoyage et de vérifier la conformité au format XML, nous avons utilisé l'éditeur Oxygen XML Editor²⁶. Cet outil très simple à manipuler nous aide à vérifier la conformité du texte avec sa version initiale et également à valider la structure formelle de chaque texte. En effet, un document XML validé reste avant tout un XML bien formé et doit être conforme aux règles du schéma qui définit la structure et les éléments légaux d'un document XML. Dans notre corpus, ces règles sont précisées dans le schéma *teillite.rnc*. Un document XML bien construit est dans tous les cas exempt de toute erreur syntaxique. En effet, les balises hiérarchisées reflètent la structure logique du document (sections, chapitres, parties, préface, introduction, annexe, etc.), les segments textuels spécifiques (titres, sous-titres, énumérations, appels de note, citations, etc.) sont identifiés.

Expliquons plus en détail l'annotation d'ordre structural dans notre corpus. Notre schéma respecte les recommandations P5 de la TEI et comprend l'annotation de l'en-tête et l'annotation du texte (Falaise, 2014). L'en-tête du corpus différencie l'en-tête du texte en ce sens que le corpus *header* permet de noter ce qui s'applique à tous les textes, le header de chaque texte ne notant que ses spécificités ou exceptions par rapport à ce qui est défini pour

²² HTML HyperText MarkupLanguage

²³ Perl un langage semi-interprété <http://www.perl.org/>

²⁴ Nous tenons à remercier tout particulièrement Marie-Paule Jacques de nous avoir aidée à traiter tout le corpus sous Perl.

²⁵ XML Extensible Markup Language

²⁶ Oxygen xml editor : <http://www.oxygenxml.com/>

l'ensemble du corpus. L'en-tête du texte comporte 4 composants dont <fileDesc>, <encodingDesc>, <profileDesc>, <revisionDesc>.

```

<teiCorpus>
  <teiHeader>en-tête du corpus</teiHeader>
  <TEI>
    <teiHeader>en-tête du premier texte</teiHeader>
      <fileDesc></fileDesc>
      <encodingDesc></encodingDesc>
      <profileDesc></profileDesc>
      <revisionDesc></revisionDesc>
    <text>
      <front>textes liminaires : en-têtes, pages de titre, préface</front>
      <body>corps du texte
        <div1>section de niveau 1 (sections, chapitres, parties)
          <p>paragraphe <s>phrase ou syntagme</s><s>...</s></p>
        </div1>
        <div2>section de niveau 2 (sections, chapitres, parties)
          <p><s>phrase ou syntagme</s></p><p>...</p>
        </div2>
      </body>
      <back>annexes, bibliographies, index</back>
    </text>
  </TEI><TEI>deuxième texte...</TEI>...
</teiCorpus>

```

Figure 3-1 : Structure d'un document TEI Lite

La Figure 3-1 ci-dessus présente ci-dessus la structure arborescente du schéma Teillite (TEI P5). La balise <fileDesc> décrit la bibliographie complète du document (titre, auteur, éditeur, conditions juridiques de distribution, etc.), <encodingDesc> décrit la relation entre le texte balisé et ses sources, c'est-à-dire les stratégies d'encodage adoptées, les niveaux d'encodage, les balises utilisées, <profileDesc> apporte des renseignements sur le contexte de constitution du corpus (sujet, langues, situation de production, individus participants, etc.), <revisionDesc> détaille les révisions subies par le corpus, en quelques sorte, le livre de bord des responsables du corpus.

L'annotation de la structure du texte se compose de trois grandes parties : <front>, <body>, <back>. Dans <front> figurent toutes les informations qui portent sur le titre, les auteurs, l'adresse, l'affiliation, le résumé, les mots-clés <head>, <docAuthor>, <address>. Dans le <body>, on trouve la plupart des informations, y compris les titres de section, les sections, les figures. Toutes les informations sur les références, les remerciements, les notes, les annexes, etc. sont recensées dans la section <back>. Nous soulignons l'importance du processus d'analyse linguistique et d'annotation du corpus à la fois sur le plan linguistique et

didactique. En effet, cette annotation du corpus permet d'une part d'enrichir les ressources lexicales, pour ensuite effectuer des recherches assez fines et d'autre part de favoriser la mise en place de différentes séquences didactiques. Tutin (2010) insiste également sur le rôle de l'annotation linguistique et sémantique dans le fait que cette technique est à même de « générer des bases de données lexicales, comportant des informations d'usage, mais aussi une navigation entre données lexicales et exemples textuels, et d'imaginer des parcours textuels entre collocations, sur la base de l'analogie sémantique et syntaxique » (*ibid.* : 143).

Toutes les traces de l'auteur (y compris les en-têtes, les pieds de page, des notes, des résumés), sont respectées, toutes les balises logiques sont bien vérifiées tout en suivant toutes les recommandations adoptées par la TEI²⁷. Oxygen nous guide dans la création des balises, et dans la vérification des segments textuels par <div>. Le nettoyage représente un procédé fastidieux. À l'issue de cette phase d'annotation, notre corpus en SHS sera intégré au corpus Scientext en ligne et prêt pour les requêtes.

3.1.6. Droit d'exploitation et de reproduction des textes

Une partie de notre corpus les sous-corpus de quatre disciplines (linguistique, psychologie, économie, sciences du langage) est diffusable étant donné que les auteurs ont signé un contrat d'acceptation de publication de leur travail dans le cadre du projet Scientext²⁸. Les autres articles du corpus ne sont pas encore publiés du fait du droit d'auteur et sont consultables en intranet pour notre recherche scientifique. La liste des textes retenus se trouve en *Annexe 2 : Liste des revues*. Les textes libres de droit sont interrogeables en ligne sur : <http://scientext.msh-alpes.fr/> (Falaise, 2014).

3.3. Repérage des séquences polylexicales

Les séquences polylexicales de la phraséologie transdisciplinaire, quel que soit leur degré de figement ou leur structure syntaxique, sont partagées par les disciplines et servent à décrire les activités et le raisonnement scientifiques. Ces expressions sont souvent des *segments répétés*, c'est-à-dire, des suites de mots récurrentes ou selon Salem (1987), « les segments dont la fréquence est supérieure ou égale à 2 » (*ibid.* : 50). Il existe des segments

²⁷ TEI TextEncoding Initiative : <http://www.tei-c.org/index.xml>

²⁸ Rapport du projet Scientext, Annexe 1 : Contrats établis avec les auteurs pour l'utilisation du texte en ligne et sa diffusion éventuelle. Consulté le 22 octobre 2014 : <http://scientext.msh-alpes.fr/scientext-site/spip.php?article14>

répétés continus et discontinus. Les segments répétés continus comprennent toujours les mêmes items dans le même ordre. Par exemple, *il en résulte qu'* est considéré comme un segment répété continu, mais *il... que* est un segment répété discontinu, car de nombreux éléments peuvent être insérés entre *il... que* par exemple *il en résulte que, il est indiqué que*, etc. Dans notre recherche, nous nous intéressons uniquement aux segments répétés continus qui correspondent aux expressions toutes faites comme *en d'autres termes, par contre*, etc.

Pour traiter les séquences polylexicales, plusieurs types d'extraction ont été mis en œuvre afin de repérer les différents types d'expressions. L'approche distributionnelle propose notamment deux méthodes d'extraction, dont les N-grammes et la co-occurrence (Granger & Paquot, 2008) (Figure 3-2).

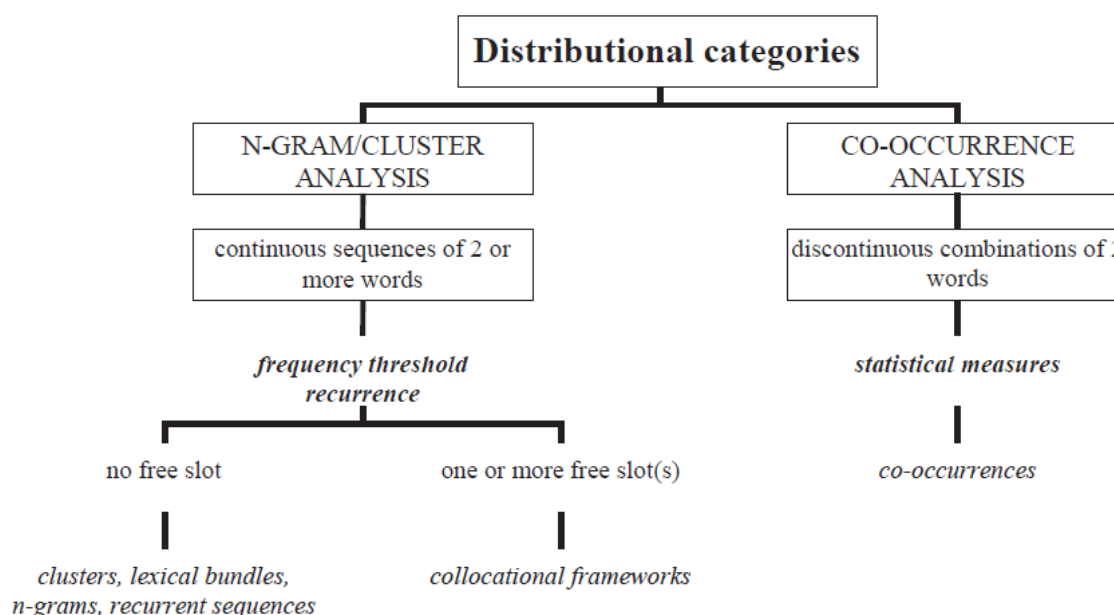


Figure 3-2 : Distributionnelle catégorie (d'après Granger & Paquot, 2008 : 39)

D'un côté, l'analyse des co-occurrences ou *co-occurrence analysis* prend appui sur les associations statistiquement récurrentes des unités lexicales, mais qui ne sont pas nécessairement immédiatement contiguës. Les mots qui apparaissent souvent ensemble sont considérés comme des collocations (Sinclair, 1991 ; Stubbs, 2002). Les collocations sont définies comme des cooccurrences habituelles afin d'identifier des séquences disjointes de deux mots entretenant une relation syntaxique et apparaissant souvent ensemble dans un petit contexte. Cette procédure nous permet de récupérer des groupes de mots, leur fréquence et l'attraction des items dans une paire, mais n'est pas particulièrement adaptée au repérage des

séquences avec des éléments contigus. Ces techniques permettent d'analyser les collocations comme cela a été proposé dans l'équipe du LIDILEM dans le cadre du projet EMOLEX²⁹.

D'un autre côté, la technique des N-grammes ou *N-Gram/Cluster analysis* consiste en une méthode qui permet d'extraire des segments répétés, des combinaisons grammaticalisées et de repérer des séquences linéaires de deux ou plusieurs mots. De fait, cette méthode ne prend en compte ni l'idiomaticité ni le statut structural des expressions. C'est pourquoi elle intéresse des auteurs ayant différents objectifs (extraction terminologique, études interlangues, etc.). Les séquences extraites reçoivent plusieurs appellations : *segments répétés* (Salem, 1987), *N_grammes* (Stubbs, 2002), *lexical bundles* (Biber, Conrad & Cortes, 2004 b ; Biber & Conrad, 1999 ; Biber *et al.*, 2003), et *séquences récurrentes* (De Cock, 2004). Cette dernière technique, très simple, est tout à fait suffisante pour notre objectif.

3.2.1. Extraction

Dans notre recherche, nous adoptons seulement la méthode des n-grammes, car nous travaillons sur des séquences d'éléments contigus, et avons utilisé un programme créé par Olivier Kraif du LIDILEM dont la fonctionnalité est semblable à la boîte à outils N_grammes Statistical Package NSP pour repérer les segments répétés. Nous utilisons par ailleurs les résultats de l'analyseur syntaxique de corpus Syntex, en repérant les locutions déjà identifiées par le programme. Nous allons maintenant préciser la procédure d'extraction des séquences polylexicales dans notre corpus.

3.2.1.1. Des segments répétés

Les segments répétés sont recensés par le logiciel sur une simple base statistique dans la mesure où ils apparaissent régulièrement dans une fenêtre de *x mots*. Dans notre cas, nous avons fixé la taille minimale de la fenêtre à 2 mots et la taille maximale à 7 mots. Il est à noter que notre corpus a été lemmatisé et que la recherche des segments répétés se fait sur la suite de lemmes, discipline par discipline. Le logiciel génère à la fin pour chaque discipline une liste des segments répétés classés selon leur longueur, avec la fréquence et la catégorie syntaxique des composants, sachant que les segments répétés et les locutions sont repérés par leur lemme afin de regrouper toutes les formes fléchies.

Nous avons recouru au système de gestion de base de données Access dans le but de confronter les résultats de chaque discipline et de rendre compte de la répartition des

²⁹ Site du projet Emolex : <http://www.emolex.eu/>

segments répétés sur tout le corpus. Nous avons obtenu un résultat assez important de segments répétés sous forme lemmatisée de longueur variée (45 184 segments répétés au total, dont 40 733 d'une longueur de 2 mots). Cette procédure génère beaucoup de bruit, car beaucoup d'associations ne constituent pas des expressions. Il s'agit de suites de mots qui ne correspondent pas à un constituant par exemple *ce être le*³⁰ (1001 occurrences, 10 disciplines), *que il ne* (356 occurrences, 10 disciplines), *il ne être pas* (312 occurrences, 10 disciplines) ou n'ont pas de valeur sémantique, comme le cas de *sur le* (1142 occurrences, 10 disciplines). Nous avons enlevé des groupements qui restent neutres (*il y avoir* (938 occurrences sur 10 disciplines), *ce être le cas* (180 occurrences, 10 disciplines), etc.) ou ceux qui ne nous intéressent pas sur le plan didactique (*presse universitaire* (171 occurrences, sur 10 disciplines), *comme le processus de* (410 occurrences sur 10 disciplines), etc.). Tous les regroupements relativement généraux ou difficiles à extraire du contexte sont également à exclure, par exemple les associations nominales de type *l'hypothèse de* (300 occurrences sur 10 disciplines), *l'idée de* (104 occurrences, 9 disciplines), *le choix de* (48 occurrences, 5 disciplines), *l'importance de* (274 occurrences, 10 disciplines), etc. ainsi que des structures verbales (*se appuyer sur* (261 occurrences, 10 disciplines), *être considérer comme* (249 occurrences, 10 disciplines), etc.). Le débroussaillage se fait manuellement et s'avère fastidieux, mais reste le seul moyen pour relever les associations susceptibles d'être intéressantes pour notre travail. Cependant, le nombre de segments répétés repérés à cette étape reste minime par rapport à ce que nous avons envisagé. En effet, un nombre très élevé des groupes de mots avait été analysé automatiquement comme locutions par Syntex et celles-ci sont récupérées uniquement par des requêtes en ligne sur l'interface du corpus Scientext.

3.2.1.2. Extraction des locutions dans les corpus analysés par Syntex

Ces extractions ont été réalisées à l'aide de l'interface Scientext (Figure 3-3). Dans une première étape, on sélectionne dans l'interface différents sous-corpus correspondant aux différentes disciplines. Grâce à l'annotation effectuée, l'utilisateur peut sélectionner la discipline, le type de documents et les parties du texte sur lesquels il désire travailler.

³⁰ Tous les exemples sont repérés par leurs lemmes dans le corpus.

1. Choix des types de textes		Affiner la sélection: Suite	
Ici, vous pouvez sélectionner un sous-ensemble du corpus, en fonction des disciplines, des genres et des parties textuelles.			
291 textes sont sélectionnés (2 698 667 mots) [Voir les détails]			
Accueil Choix des textes ► 1. Types de textes 2. Liste de textes Recherche dans les textes 3. Mode de recherche 4. Recherche Résultats 5. Concordancier 6. Statistiques Langue de l'interface: Français ▼ Mode d'emploi	Disciplines <input type="checkbox"/> Sciences humaines <input checked="" type="checkbox"/> Linguistique <input checked="" type="checkbox"/> Psychologie <input checked="" type="checkbox"/> Sciences de l'éducation <input type="checkbox"/> Traitement Automatique des Langues <input checked="" type="checkbox"/> Économie <input checked="" type="checkbox"/> Sciences politiques <input checked="" type="checkbox"/> Anthropologie <input checked="" type="checkbox"/> Histoire <input checked="" type="checkbox"/> Géographie <input checked="" type="checkbox"/> Sciences de l'information et de la communication <input checked="" type="checkbox"/> Sociologie <input type="checkbox"/> Sciences expérimentales <input type="checkbox"/> Biologie <input type="checkbox"/> Médecine <input type="checkbox"/> Sciences appliquées <input type="checkbox"/> Électronique <input type="checkbox"/> Mécanique	Types de documents <input checked="" type="checkbox"/> Article <input type="checkbox"/> Communication <input type="checkbox"/> Thèse <input type="checkbox"/> HDR	Parties <input checked="" type="checkbox"/> Parties principales <input checked="" type="checkbox"/> Développement <input checked="" type="checkbox"/> Introduction <input checked="" type="checkbox"/> Conclusion <input checked="" type="checkbox"/> Autres parties <input checked="" type="checkbox"/> Résumé <input checked="" type="checkbox"/> Notes <input checked="" type="checkbox"/> Titres <input checked="" type="checkbox"/> Remerciements <input checked="" type="checkbox"/> Annexe

Figure 3-3 : Liste des disciplines dans le corpus SHS – Interface Scientext (corpus Intranet)

L'exploitation du corpus se déroule en trois temps : la sélection du corpus, la recherche dans les textes, l'affichage et le traitement des résultats. La Figure 3-3 présente notre corpus de travail composé d'articles de différentes disciplines en SHS (linguistique, psychologie, sciences de l'éducation, économie, sciences politiques, anthropologie, histoire, géographie, sciences de l'information et de la communication, sociologie). Dans notre cas, nous nous intéressons au fonctionnement des séquences polylexicales dans toutes les parties textuelles de ces articles scientifiques.

63 textes sont sélectionnés (413 915 mots) [Voir les détails]			
Choix des textes ► 1. Types de textes 2. Liste de textes Recherche dans les textes 3. Mode de recherche 4. Recherche Résultats 5. Concordancier 6. Statistiques Langue de l'interface: Français ▼ Mode d'emploi Compte utilisateur Connexion	Disciplines <input type="checkbox"/> Sciences humaines <input checked="" type="checkbox"/> Linguistique <input type="checkbox"/> Psychologie <input type="checkbox"/> Sciences de l'éducation <input type="checkbox"/> Traitement Automatique des Langues <input type="checkbox"/> Économie <input type="checkbox"/> Sciences politiques <input type="checkbox"/> Anthropologie <input type="checkbox"/> Histoire <input type="checkbox"/> Géographie <input type="checkbox"/> Sciences de l'information et de la communication <input type="checkbox"/> Sociologie <input type="checkbox"/> Sciences expérimentales <input type="checkbox"/> Biologie <input type="checkbox"/> Médecine <input type="checkbox"/> Sciences appliquées <input type="checkbox"/> Électronique <input type="checkbox"/> Mécanique	Types de documents <input checked="" type="checkbox"/> Article <input type="checkbox"/> Communication <input type="checkbox"/> Thèse <input type="checkbox"/> HDR	Parties <input checked="" type="checkbox"/> Parties principales <input checked="" type="checkbox"/> Développement <input checked="" type="checkbox"/> Introduction <input checked="" type="checkbox"/> Conclusion <input checked="" type="checkbox"/> Autres parties <input checked="" type="checkbox"/> Résumé <input checked="" type="checkbox"/> Notes <input checked="" type="checkbox"/> Titres <input checked="" type="checkbox"/> Remerciements <input checked="" type="checkbox"/> Annexe
	Tout Rien	Tout Rien	Tout Rien
Sauvegarder la sélection: [TEXTS] Restaurer une sélection: [sélection-linguistique.txt] Parcourir... OK			

Figure 3-4 : Sélectionner des articles en linguistique

Choix des textes	30 textes sont sélectionnés (147 255 mots) [Voir les détails]	Textes triés par Discipline - 10 textes par page.
1. Types de textes		
2. Liste de textes	<input type="checkbox"/> Peut-on parler de sémiographie optimale ? (Jaffré) <i>LIDIL 30 - Acquisition et enseignement de la morphographie</i> Linguistique / Article [lin-art-2]	
Recherche dans les textes	<input checked="" type="checkbox"/> Langage intérieur et ontologie linguistique à la fin du XIXe siècle. (Puech) <i>Langue Française 132, pages 26-47</i> Linguistique / Article [lin-art-8]	
3. Mode de recherche		
4. Recherche	<input type="checkbox"/> Contre la trichotomie Syntaxe/Sémantique/Pragmatique (Laurendeau) <i>Revue de Sémantique et Pragmatique 1, pages 115-131</i> Linguistique / Article [lin-art-34]	
Résultats		
5. Concordancier	<input type="checkbox"/> Sémantique, pragmatique et missions de la linguistique (Vion) <i>Revue de Sémantique et Pragmatique 1, pages 81-93</i> Linguistique / Article [lin-art-44]	
6. Statistiques		
Langue de l'interface:	<input type="checkbox"/> Tout contre vs. très contre (Leeman) <i>Travaux de Linguistique 42-43, 241-252</i> Linguistique / Article [lin-art-63]	
Mode d'emploi	<input type="checkbox"/> Scènes d'actions radiophoniques et prises de rôles : informer, débattre, divertir (Burger) <i>Revue de Sémantique et Pragmatique 7, pages 179-196</i> Linguistique / Article [lin-art-71]	
	<input type="checkbox"/> L'emploi spatial de contre : propositions pour un traitement unifié (Dendale) <i>Travaux de Linguistique 42-43, 229-239</i> Linguistique / Article [lin-art-75]	
Compte utilisateur	<input checked="" type="checkbox"/> Repérage automatique de structures linguistiques en corpus : le cas des énoncés définitoires (Rebeyrolle, Tanguy) <i>Cahiers de Grammaire 25</i> Linguistique / Article [lin-art-83]	
Connexion	<input checked="" type="checkbox"/> Sens et contre sens : du cotexte au contexte (Cusin-Berche) <i>Revue de Sémantique et Pragmatique 1, pages 135-146</i> Linguistique / Article [lin-art-95]	

Figure 3-5 : Liste des articles sélectionnés en linguistique

Scientext nous permet d'établir et de sauvegarder une présélection d'articles dans chaque discipline pour d'autres interrogations du même corpus lors de sessions ultérieures. Dans la Figure 3-4, nous avons sélectionné le sous-corpus en linguistique ainsi que les articles sur lesquels nous voulons travailler (Figure 3-5). Après avoir sélectionné le corpus à exploiter, l'utilisateur peut accéder au contenu des textes par trois modes de recherche mis à sa disposition : le mode sémantique guidé, le mode libre guidé, et le mode avancé pour ceux qui maîtrisent la syntaxe informatique des expressions régulières. Tous ces modes sont traduits dans le même langage de requête ConcQuest (Kraif, 2008) et développés par Falaise (Falaise *et al.*, 2011). ConcQuest est un logiciel d'extraction de concordances permettant de rechercher des expressions complexes à travers un corpus de texte. Pour simplifier le prototype d'interface antérieurement développé pour ConcQuest, un formulaire spécifique de requêtes a été mis en œuvre (Falaise *et al.*, 2012). (Figure 3-6)

3. Mode de recherche	Effectuer une recherche: <input type="button" value="Suite"/>
Ici, vous pouvez choisir le mode de recherche.	
<input type="radio"/> Recherche sémantique C'est le mode le plus simple. Il permet d'effectuer une recherche sémantique à l'aide des grammaires créées par l'équipe.	
<input checked="" type="radio"/> Recherche libre Dans ce mode il est possible de chercher librement des mots, notamment en fonction de leur forme, de leur lemme, de leur catégorie, et de leurs relations syntaxiques	
<input type="radio"/> Recherche avancée C'est le mode le plus complexe. Dans ce mode il est possible de créer des grammaires, plus puissantes que la recherche libre	

Figure 3-6 : Différents modes de recherche dans Scientext

En ce qui concerne les séquences polylexicales, toutes les requêtes sont donc effectuées en mode libre guidé. Ce mode de recherche nous permet de définir nous-mêmes

des requêtes par la forme, par le lemme, par des catégories ou encore par des relations syntaxiques.

Une fois les requêtes terminées, l’affichage des résultats est variable selon les attentes de l’utilisateur, par exemple, l’affichage des concordances dans un concordancier KWIC (Key Word In Context) dont les contextes gauche et droit sont paramétrables (l’utilisateur peut visualiser jusqu’à 200 mots dans chaque contexte) (Figure 3-7). Grâce au concordancier et au contexte élargi, nous avons la possibilité de filtrer les résultats (à l’aide d’une case décochable dans chaque ligne du concordancier) afin de conserver ceux qui sont pertinents pour notre recherche. Toutes les occurrences sélectionnées peuvent être sauvegardées et exportées dans un tableur (format HTML ou CSV) pour servir au travail sémantique d’analyse des séquences polylexicales et aux illustrations didactiques lors de notre étape d’expérimentation.

5. Concordancier					Voir les statistiques: Suite
Ici, vous pouvez consulter les occurrences trouvées lors de la recherche sous forme de concordancier.					
1065 occurrences.					Page: 1 > ▶
▼ N°	▼ Contexte gauche: 10 mots	▼ Occurrence:	▼ Contexte droit: 10 mots	▼ Réf. texte	
<input checked="" type="checkbox"/> 1	D'abord , l' expression " langage intérieur " semble désigner	avant tout	pour la tradition philosophique une conception particulière non pas tant du	[lin-art-8-body]	
<input checked="" type="checkbox"/> 2	Dans cette tradition , c' est	en effet	de la pensée qu' il est dit qu' elle est un	[lin-art-8-body]	
<input checked="" type="checkbox"/> 3	Dans cette tradition , c' est	en effet	de la pensée qu' il est dit qu' elle est un	[lin-art-8-body]	
<input checked="" type="checkbox"/> 4	C' est	sans doute	dans cette forme chiasmatique qu' elle trouve à la fois ses ressources	[lin-art-8-body]	
<input checked="" type="checkbox"/> 5	est sans doute dans cette forme chiasmatique qu' elle trouve	à la fois	ses ressources et ses limites ; et l' on ne	[lin-art-8-body]	
<input checked="" type="checkbox"/> 6	et ses limites ; et l' on ne doit	sans doute	pas s' étonner si elle opère une résurgence au cours de l'	[lin-art-8-body]	
<input checked="" type="checkbox"/> 7	C' est	sans doute	dans cette forme chiasmatique qu' elle trouve à la fois ses ressources	[lin-art-8-body]	
<input checked="" type="checkbox"/> 8	est sans doute dans cette forme chiasmatique qu' elle trouve	à la fois	ses ressources et ses limites ; et l' on ne	[lin-art-8-body]	
<input checked="" type="checkbox"/> 9	et ses limites ; et l' on ne doit	sans doute	pas s' étonner si elle opère une résurgence au cours de l'	[lin-art-8-body]	
<input checked="" type="checkbox"/> 10	: cette " permanence " (?) atteste	au moins	de la stabilité d' une dénomination (néanmoins variable)	[lin-art-8-body]	
<input checked="" type="checkbox"/> 11	1.1.4	En effet	, au Moyen -Âge , le thème du langage intérieur	[lin-art-8-body]	
<input checked="" type="checkbox"/> 12	du support matériel de la pensée , c' est ,	d' autre part	, l' assignation d' un degré d' intériorité ou d'	[lin-art-8-body]	
<input checked="" type="checkbox"/> 13		Dans tous les cas	, l' universalité de la faculté de langage doit affronter	[lin-art-8-body]	
<input checked="" type="checkbox"/> 14	Il dégage	pour son compte	trois points de convergence qui peuvent nous apparaître comme autant	[lin-art-8-body]	
<input checked="" type="checkbox"/> 15	1 3 .	D' autre part	, la thématique de la pensée comme " dialogue de	[lin-art-8-body]	
<input checked="" type="checkbox"/> 16	de l' âme avec elle - même " implique	par essence	une distinction entre un discours intérieur purement conceptuel (même si	[lin-art-8-body]	
<input checked="" type="checkbox"/> 17	créature et créateur (dans la prière silencieuse ,	par exemple	, à celle psychologico-linguistique du sujet parlant et de	[lin-art-8-body]	
<input checked="" type="checkbox"/> 18	" l' habitude " et l' apprentissage (phonatoire ,	par exemple	, le dédoublement (normal et pathologique) du	[lin-art-8-body]	
<input checked="" type="checkbox"/> 19	des mots n' est plus qu' un cas (à peine) particulier de celui de la représentation des objets sensibles	[lin-art-8-body]	

Figure 3-7 : Visualisation en concordancier des occurrences des locutions adverbiales en Linguistique

À propos des statistiques des résultats, Scientext permet à l’utilisateur d’extraire une liste des lemmes correspondant à chaque requête (Figure 3-8) ou de les répartir en fonction de différents critères comme les disciplines, les genres textuels, les parties textuelles (Figure 3-9). Dans la Figure 3-8, s’affichent les résultats de la requête des locutions adverbiales en linguistique dans l’ordre décroissant des occurrences. *Par exemple, en effet* restent les séquences polylexicales les plus utilisées dans cette discipline au total 115 et 109 occurrences respectivement (c’est-à-dire 10,80 % et 10,20 % sur le sous-corpus linguistique).

Dans la Figure 3-9, les résultats sont regroupés par parties textuelles. Nous constatons que les locutions adverbiales sont essentiellement utilisées dans la partie « Développement » en linguistique.

6. Statistiques
Ici, vous pouvez consulter les statistiques concernant les occurrences trouvées lors de la recherche.

Au total 1065 occurrences ont été trouvées.

Quelles statistiques souhaitez-vous consulter ?

Liste des lemmes ▼

▼ Lemme	▼ Nombre	Forme
/par/ /exemple/	115 (10.80%)	par exemple (88) , Par exemple (27)
/en/ /effet/	109 (10.23%)	en effet (58) , En effet (51)
/à/ /la/ /fois/	39 (3.66%)	à la fois (39)
/d'autre/ /part/	39 (3.66%)	d'autre part (30) , D'autre part (9)
/en/ /particulier/	36 (3.38%)	en particulier (35) , En particulier (1)
/au/ /moins/	33 (3.10%)	au moins (33)
/par/ /contre/	28 (2.63%)	Par contre (14) , par contre (14)
/sans/ /doute/	23 (2.16%)	sans doute (23)
/en/ /revanche/	22 (2.07%)	en revanche (12) , En revanche (10)
/a/ /priori/	21 (1.97%)	a priori (20) , A priori (1)
/par/ /ailleurs/	21 (1.97%)	Par ailleurs (16) , par ailleurs (5)
/en/ /d'autres/ /termes/	17 (1.60%)	En d'autres termes (14) , en d'autres termes (3)
/en/ /partie/	16 (1.50%)	en partie (16)
/de/ /fait/	16 (1.50%)	De fait (11) , de fait (5)
/tout/ /d'abord/	15 (1.41%)	tout d'abord (11) , Tout d'abord (4)
/dès/ /lors/	12 (1.13%)	dès lors (6) , Dès lors (6)
/en/ /quelque/ /sorte/	11 (1.03%)	en quelque sorte (11)

Figure 3-8 : Visualisation en lemme des résultats des locutions adverbiales en linguistique

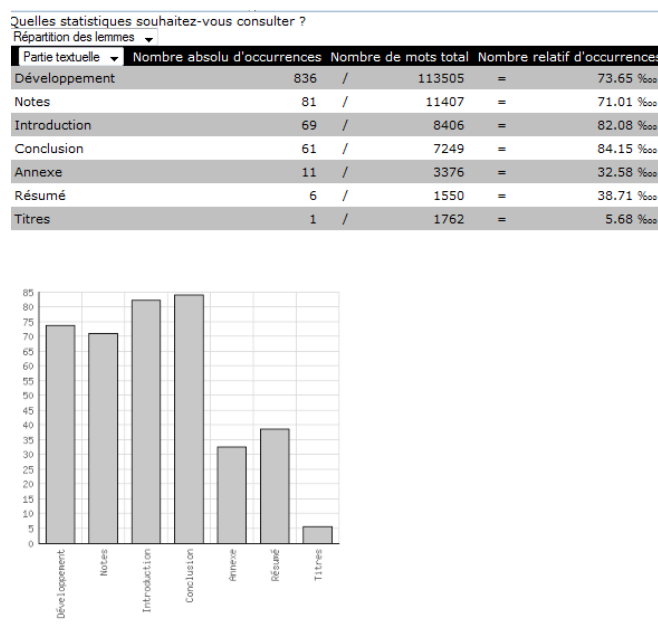


Figure 3-9 : Répartition en partie textuelle des résultats des locutions adverbiales

Comme l'extraction des segments répétés, celle des locutions est également effectuée discipline par discipline pour chaque fonctionnement syntaxique et les résultats sont

sauegardés dans un fichier Excel. Tous les lemmes, leurs formes³¹, leur fréquence absolue, leur fréquence relative sont pris en considération. Leurs formes nous semblent intéressantes sur le plan didactique lors de la conception des activités pédagogiques. La fréquence absolue générée automatiquement définit le pourcentage d'occurrences de chaque segment répété sur le nombre total d'occurrences pour chaque requête, tandis que la fréquence relative résulte du pourcentage du nombre d'occurrences du segment répété sur le nombre de mots du corpus correspondant à la recherche. Les fréquences absolue et relative sont indispensables pour les analyses quantitatives. Afin de confronter les résultats de différentes disciplines pour une même requête, nous avons regroupé dans un fichier tous les lemmes partageant la même catégorie syntaxique et enlevé tous les doublons. Les dix fichiers-disciplines ainsi que ce nouveau fichier sont ensuite importés dans Access et interrogés dans la base de données. Le résultat final comprend une liste des lemmes dans les dix disciplines, leurs formes, leur fréquence, ainsi que leur répartition (Figure 3-10). Cela sera ensuite exporté dans un tableur Excel.

	A	B	C	D	E	F	G	H	I	J	AF	AG	AH
	Lemme	linguistique_F	linguistique_Nc	linguistique_F	psychologie_F	psychologie_Nc	psychologie_F	education_Foi	education_Nor	Education_Rd	C1	C2	
5	/même/ /si/	même si (25), Même si (7)	32	2,17310108315 507	même si (25), Même si (4)	29	1,44506510267 437	même si (27), Même si (14)	41	2,10061430159 698	424	10	
6	/bien/ /que/	Bien que (7), bien qu' (7), bien que (6), Bien qu' (1)	21	1,42609758582 052	Bien que (14), Bien que (12), bien qu' (7), Bien qu' (3)	36	1,79387392056 128	bien qu' (14), bien qu' (8), Bien que (6), Bien qu' (2)	30	1,53703485482 706	388	10	
7	/tandis/ /que/	tandis que (5), tandis qu' (1)	6	0,40745645309 1576	tandis que (12), tandis qu' (5), Tandis que (2)	19	0,94676679140 7344	tandis que (22), tandis qu' (5), Tandis que (3)	30	1,53703485482 706	349	10	
8	/ou/ /encore/	ou encore (19), Ou encore (2)	21	1,42609758582 052	ou encore (25), Ou encore (1)	26	1,29557560929 426	ou encore (17)	17	0,87098641773 5333	346	10	
9	/dans/ /la/ /mesure/ /où/	dans la mesure où (28), Dans la mesure où (8)	36	2,44473871854 945	dans la mesure où (23), Dans la mesure où (1)	24	1,19591594704 086	dans la mesure où (17), Dans la mesure où (1)	18	0,92222091289 6235	255	10	
10	/pour/ /que/	pour que (11), pour qu' (2)	13	0,88282231503 1748	pour que (11), Pour que (1)	12	0,59795797352 0428	pour que (13), pour qu' (9), Pour qu' (3), Pour que (3)	28	1,43456586450 525	255	10	
11	/ou/ /bien/	ou bien (23)	23	1,56191640351 771	ou bien (9)	9	0,44846848014 0321	ou bien (10), Ou bien (2)	12	0,61481394193 0823	161	10	
	plutôt que (4),			0,54327527078	plutôt que (13)		0,99659662253	plutôt que (11)		0,92222091289			

Figure 3-10 : Confrontation des résultats dans Access pour des locutions adverbiales

Sur la Figure 3-10, tous les lemmes qui ont la même structure syntaxique sont recensés dans la colonne « Lemme ». Chaque discipline est représentée par trois valeurs : les formes, la fréquence absolue (le nombre d'occurrences), et leur fréquence relative. Le nombre d'occurrences de chaque séquence polylexicale sur tout le corpus est calculé par le logiciel et affiché à la colonne C1. La colonne C2 représente le nombre des disciplines où apparaît le

³¹ Ou formes fléchies, formes conjuguées ou accordées d'un mot de base

lemme. Par exemple, */même/ /si/* qui apparaît dans 10 disciplines et fait un total de 424 occurrences, alors que */dans/ /la/ /mesure/ /où/* apparaît 255 fois dans les 10 disciplines. Nous avons également la possibilité de filtrer notre résultat en fonction de notre objectif de recherche. En effet, la double tâche d'extraction nous a permis d'obtenir des résultats intéressants. Pour pouvoir mettre en évidence notre phraséologie transdisciplinaire, il nous faut nous baser sur des critères précis et rigoureux dans la sélection des séquences polylexicales.

3.2.2. Critères de sélection

Les séquences polylexicales retenues doivent répondre à différents critères à savoir la fréquence, la répartition disciplinaire et la discursivité. Les méthodes statistiques de fréquence et de répartition restent les plus utilisées pour extraire des groupes composés et des collocations (Phal, 1971 ; Coxhead, 2002 ; Tutin, 2008, 2014 ; Drouin, 2007 ; Da Sylva, 2010 ; Paquot & Bestgen, 2009 ; Paquot, 2010). Da Sylva (2010) envisage de concevoir des ressources lexicales composées du vocabulaire servant de base et du vocabulaire spécialisé pour constituer des entrées d'index structurées. L'expérimentation d'extraction semi-automatique des ressources lexicales se fait à partir de deux corpus : l'un pour l'anglais (14 millions de mots) et l'autre pour le français (2,5 millions de mots). Les fréquences d'occurrences des noms lemmatisés ont été comptées afin de récupérer les mots-clés les plus fréquents.

Paquot & Bestgen (2009) proposent d'extraire des mots-clés du discours académique en comparant le sous-corpus académique avec un sous-corpus de fiction de BNC British National Corpus (BNC). Ils se basent sur trois méthodes statistiques : loge-likelihood ratio, t-test et Wilcoxon-Mann-Whitney. Les trois listes issues des trois méthodes d'extraction sont ensuite comparées sur plusieurs critères, à savoir le nombre de mots-clés extraits à chaque méthode, leur ratio et leur répartition. Cette confrontation permet d'identifier la méthode statistique la plus avantageuse pour extraire des mots-clés. En ce qui nous concerne, nous choisirons une méthode plus simple, tenant principalement compte de l'aspect transdisciplinaire.

3.2.2.1. Fréquence

Comme nous venons de l'expliquer, les requêtes sur le corpus nous permettent d'obtenir le résultat final qui recense tous les lemmes sur tout notre corpus SHS, ainsi que

leurs formes, leur fréquence et leur répartition. Par le critère de « fréquence », nous entendons un nombre d'occurrences suffisamment important afin de répondre aux critères de *représentativité*. Pour des raisons de simplicité et de mise en œuvre, nous nous intéressons uniquement à la fréquence absolue, c'est-à-dire le nombre total d'occurrences d'un mot donné dans le corpus, et pas à la fréquence relative, soit le rapport du nombre d'occurrences d'un mot au nombre de mots que compte le corpus dans lequel le mot est apparu.

Nous avons retenu uniquement les séquences polylexicales qui apparaissent au moins **7 fois** sur tout notre corpus. Dans le Tableau 3.2 ci-dessous, nous avons relevé d'une manière aléatoire 12 séquences polylexicales de catégorie syntaxique différente et leur fréquence. En nous appuyant uniquement sur le critère de fréquence, toutes les séquences lexicales dans le tableau sont retenues, sauf à la condition que. Cependant, le seul critère de fréquence peut susciter quelques réserves, car on peut obtenir de nombreuses séquences polylexicales dont la fréquence est élevée, mais qui n'apparaissent pas dans différentes disciplines : il s'agit dans ce cas de la terminologie de la discipline. Par exemple, nous retrouvons 51 occurrences de *discrimination sur âge* uniquement en sociologie, 60 occurrences pour *culture d'information* en SIC, etc., c'est pourquoi nous proposons d'y ajouter le deuxième critère de répartition disciplinaire.

3.2.2.2. Répartition disciplinaire

Le seuil du nombre d'occurrences à prendre en compte dépend de l'application et du corpus. Pour notre part, nous avons décidé de conserver uniquement les séquences polylexicales apparaissant dans **4 disciplines** au moins.

Le critère de répartition revêt un rôle essentiel en établissant la liste de la phraséologie transdisciplinaire de base. La répartition précise le nombre de disciplines dans lesquelles le mot a été rencontré. En effet, comme nous l'avons expliqué précédemment, si un mot a seulement une très forte fréquence dans une seule discipline, il s'agit probablement d'un mot spécialisé de cette discipline. La combinaison de deux critères (fréquence et répartition) constitue une indication intéressante pour relever les séquences polylexicales transdisciplinaires et les plus significatives des écrits scientifiques. Mackey (1971) a insisté sur la prise en compte du critère de répartition en affirmant qu'« il n'y a pas que le nombre de fois où un élément est repéré qui soit important, il y a aussi la distribution ou les zones de concentration des répétitions dans l'ensemble du cours » (Mackey, 1971, cité par Chetouani, 1988 : 81).

Au Tableau 3.2, les séquences polylexicales qui répondent aux 2 critères de fréquence et de répétition sont retenues et notées **1**, les séquences exclues sont notées **0**. Seul *à la condition que* n'est pas retenu, parce qu'il ne remplit pas le critère de fréquence comme nous l'avons vu. Après avoir obtenu la liste des séquences polylexicales répondant aux critères de fréquence et de répartition, il reste à savoir si elles sont toutes intéressantes pour notre étude linguistique et sémantique de la phraséologie transdisciplinaire de base.

Séquences polylexicales	Fréquence	Répartition	Résultat
1. En particulier	564	10	1
2. Compte tenu	106	10	1
3. À l'époque	87	10	1
4. Au premier abord	11	7	1
5. Au centre de	89	10	1
6. Vu que	32	9	1
7. En toute rigueur	13	4	1
8. Tout de suite	12	6	1
9. À l'abri de	11	4	1
10. Par manque de	8	5	1
11. Sous réserve de	7	4	1
12. À la condition que	6	5	0

Tableau 3.2 : Sélection des séquences polylexicales : critère de fréquence et de répartition

3.2.2.3. Discursivité

Aux critères statistiques (fréquence et répartition), s'ajoute le critère d'ordre sémantique et discursif. Ce critère sert à relever uniquement les unités polylexicales qui jouent un rôle discursif, c'est-à-dire celles qui jouent le rôle d'organisateur de texte et de structuration du discours. Contrairement aux critères précédents, celui-ci est plutôt intuitif. À titre d'illustration, nous reprenons les mêmes exemples de séquences polylexicales retenus à la suite des tests de fréquence et de répartition. Au Tableau 3.3, toutes les séquences polylexicales qui nous intéressent au plan sémantique sont notées +, les autres sont marquées - dans la colonne « Discursivité ». Par exemple, dans Tableau 3.3 ci-dessous, les séquences polylexicales *en particulier*, *compte tenu* seront conservées, car elles répondent à la fois au critère de fréquence, de répartition et au critère sémantique. D'autres séquences polylexicales comme *à l'époque*, *au centre de* sont éliminées, même si elles apparaissent de maintes reprises dans notre corpus et dans de nombreuses disciplines. En fait, ces séquences fonctionnent comme des séquences polylexicales à valeur référentielle pour évoquer l'espace et le temps.

Séquences polylexicales	Fréquence	Répartition	Discursivité
1. En particulier	564	10	+
2. Compte tenu	106	10	+
3. À l'époque	87	10	-
4. Au premier abord	11	7	+
5. Au centre de	89	10	-
6. Vu que	32	9	+
7. En toute rigueur	13	4	+
8. Tout de suite	12	6	-
9. À l'abri de	11	4	-

10. Par manque de	8	5	+
11. Sous réserve de	7	4	+

Tableau 3.3 : Sélection des séquences polylexicales : critères de fréquence, de répartition et de pertinence

Le fastidieux dépouillement manuel reste indispensable pour relever des séquences polylexicales intéressantes pour l'analyse linguistique et les applications didactiques. Notre analyse montre que de nombreuses séquences polylexicales traditionnellement abordées dans les méthodes de français ont pourtant une fréquence relativement faible dans notre corpus, par exemple *sous prétexte de* (4 occurrences, dans 3 disciplines), *sous condition de* (3 occurrences, dans 3 disciplines). En revanche, certaines séquences rarement abordées apparaissent fréquemment, par exemple *à l'instar de* (107 occurrences, 9 disciplines), *conformément à* (64 occurrences, 10 disciplines). Les séquences polylexicales éliminées du fait de leur faible fréquence pourront être examinées dans un second temps.

Nous décrivons maintenant les approches que nous allons utiliser pour des analyses linguistiques et pour les activités didactiques.

3.4. Approches adoptées

L'approche onomasiologique constitue un moyen discursif efficace qui correspond à nos objectifs linguistiques et didactiques. Cette approche permet un accès par le sens des éléments lexicaux, mais pas par leur forme comme dans l'approche sémasiologique. Sur le plan linguistique, une telle approche nous permet de regrouper les éléments synonymes ou les expressions sémantiquement proches. Dans la typologie des séquences lexicalisées à fonction discursive que nous allons mettre en place, nous nous appuierons sur l'approche onomasiologique pour classer les MD par leur fonction rhétorique, dont nous présenterons l'exploitation didactique au chapitre 6. (Cf. Chapitre 6 : **Aspect rhétorique**).

Il est à préciser que dans notre travail, en dehors du corpus central d'articles scientifiques, nous utiliserons deux autres corpus pour répondre à nos objectifs de recherche, à savoir le corpus Emolex³² et un corpus de mémoires d'apprenants. Le corpus Emolex est un

³² EMOLEX : *le lexique des émotions dans cinq langues européennes : sémantique, syntaxique et dimension discursive*, piloté par Iva Novakova du LIDILEM, Université Stendhal – Grenoble 3 et Peter Blumenthal de l'Université de Cologne, site de l'Emolex : <http://www.emolex.eu/>

corpus représentatif de la langue générale composé de romans et d'articles journalistiques. Il est utilisé comme un corpus de contraste pour faire ressortir les spécificités d'utilisation des MD polylexicaux dans notre corpus SHS (Cf. Chapitre 4).

Le deuxième corpus est un corpus de mémoires d'apprenants vietnamophones du français. Ce corpus, constitué pour notre étude didactique, nous permet de repérer les principaux problèmes dans l'utilisation des séquences polylexicales dans les productions écrites des étudiants. (Cf. Chapitre 6).

L'analyse des séquences lexicalisées à fonction discursive est effectuée dans une approche empirique afin de relever les régularités syntaxiques et sémantiques des MD. En outre, nous optons dans le chapitre 5 pour une analyse comparative des MD, en les faisant contraster les uns avec les autres dans l'objectif de faire découvrir aux apprenants les caractéristiques saillantes de chaque MD. L'analyse comparative des MD nous paraît pertinente à plusieurs titres :

1/ au plan linguistique, une telle comparaison permettra, comme nous le souhaitons, de mieux appréhender les différents emplois de chaque MD et des MD sémantiquement proches. Plus précisément, la démarche de comparaison va se dérouler de la manière suivante : d'abord une analyse détaillée des propriétés syntaxiques puis des valeurs sémantiques du MD le plus fréquent dans chaque catégorie. Les autres MD de même catégorie seront analysés un par un en comparaison avec le MD le plus représentatif du groupe.

2/ cette comparaison représente des avantages sur le plan cognitif. En effet, elle permet d'éviter de reprendre le même schéma d'analyse pour chaque MD, ce qui risquerait de lasser le lecteur. Avec cette approche comparative, nous ne souhaitons pas mettre à disposition des lecteurs une simple liste des emplois des MD, mais leur proposer de prendre du recul en analysant ces éléments linguistiques.

3/ cette analyse nous paraît pertinente pour l'enseignement des MD car elle permet de mettre en relief les similitudes ainsi que les différences entre les MD. Une telle analyse en contraste permet de relever les emplois spécifiques de chaque MD. Le fait de prendre en compte le MD le plus fréquent d'une même catégorie, pour ensuite le mettre en comparaison avec les autres MD devrait permettre de voir précisément ce que nous pourrions retenir pour les enseignants et les apprenants en termes de synonymie par exemple.

Conclusion

Dans ce chapitre, nous avons décrit, dans un premier temps, la constitution de notre corpus d'articles scientifiques en SHS. L'utilisation du corpus analysé linguistiquement par des techniques de TAL représente de nombreux avantages pour l'analyse linguistique et également pour les activités didactiques que nous souhaitons mettre en place par la suite. Sur le plan linguistique, nous pouvons effectuer différentes requêtes qui correspondent à nos questions de recherche. Sur le plan didactique, cela nous permettra de concevoir des activités et des ressources lexicographiques basées sur corpus, qui seront présentées au chapitre 6.

Dans un deuxième temps, nous avons défini les critères qui sous-tendent notre typologie des séquences lexicalisées à fonction discursive. Celles-ci doivent répondre à trois critères à savoir la fréquence, la répartition disciplinaire et la discursivité. Ces séquences polylexicales doivent jouer un rôle d'organisation dans le texte. Enfin, nous avons présenté différentes approches pour l'étude sur corpus des séquences lexicalisées à fonction discursive et pour les activités didactiques, à savoir l'approche onomasiologique, l'approche empirique pour une analyse syntactico-sémantique des MD polylexicaux et l'approche comparative des MD que nous avons choisi d'appliquer dans le chapitre 5 et le chapitre 6.

PARTIE 2 :

ANALYSES LINGUISTIQUES

Chapitre 4.

Typologie des séquences lexicalisées à fonction discursive

Dans le chapitre 1 et le chapitre 2, nous avons précisé les besoins de notre public universitaire. Nous avons souligné notre intérêt pour le lexique transdisciplinaire scientifique, et plus particulièrement pour les séquences lexicalisées à fonction discursive. Pour faire de ces unités lexicales un objet d'enseignement et d'apprentissage en tant que tel, il est nécessaire de mettre en place une méthodologie adéquate pour ce public. Les analyses syntaxiques et sémantiques apporteront des éclairages aux éléments à prendre en compte dans l'enseignement/apprentissage de ces éléments linguistiques.

Ce chapitre est consacré à l'élaboration d'une typologie des séquences lexicalisées à fonction discursive les plus représentatives des écrits scientifiques. La constitution de cette typologie nous intéresse à différents plans. Au plan linguistique, cette typologie servira à déterminer différents traits linguistiques des MD. Comme nous envisageons de recourir à l'approche onomasiologique, cette typologie paraît être utile au plan didactique en mettant à disposition des apprenants un accès par la fonction rhétorique des MD. En outre, cette typologie sera intégrée à une base de données consultable en ligne pour des activités didactiques (Cf. Chapitre 6).

Ce chapitre va être organisé de la manière suivante : nous allons d'abord expliciter les principes à prendre en compte avant d'élaborer une typologie des séquences lexicalisées à fonction discursive. Nous décrirons par la suite les marqueurs de chaque groupe. Parmi les MD retenus dans ce travail, notre intérêt porte essentiellement sur les marqueurs de reformulation et nous expliquerons enfin pourquoi nous avons mis l'accent sur ce type de marqueurs.

4.1. Principes sous-jacents de notre typologie

Il nous semble indispensable de préciser les principes sous-jacents qui nous guident dans la conception de cette typologie. Premièrement, notre objet de recherche, les séquences lexicalisées à fonction discursive, font partie de la phraséologie transdisciplinaire de base des écrits scientifiques. Ces unités lexicales ne sont pas absentes d'autres genres textuels, mais nous nous intéressons à leur fonctionnement syntaxique et sémantique dans un genre spécifique. Les séquences lexicalisées à fonction discursive retenues dans notre travail sont à la fois transdisciplinaires et fréquentes dans les écrits scientifiques. Elles sont recensées dans l'Annexe 1.

Deuxièmement, pour l'approche que nous adoptons, l'approche onomasiologique, une typologie qui prend appui sur les fonctions que remplissent des séquences lexicalisées à fonction discursive, nous semble adéquate. Ces séquences polylexicales vont être regroupées dans deux grands ensembles : la fonction métadiscursive et la fonction argumentative. Chaque fonction rassemble différents sous-groupes et nous en distinguons neuf.

Le premier grand groupe renvoie à tous les marqueurs qui jouent une fonction métadiscursive. Cette catégorie permet de rassembler tous les marqueurs qui ont un rôle dans l'organisation textuelle. Ils comprennent tous les marqueurs de guidage du lecteur, type de marqueurs particulièrement développé dans les écrits scientifiques. Par l'utilisation des MD, le locuteur manifeste son intention de faciliter le décodage du lecteur. Les marqueurs métadiscursifs englobent toutes les opérations de reformulation, de comparaison, d'exemplification, etc. Les marqueurs de cette catégorie font partie de la structure globale du texte. Ils montrent comment le texte s'organise ou bien font référence au texte lui-même. Ils permettent de placer une introduction (*en guise d'introduction*, etc.), une progression dans les idées (*en outre*, *par ailleurs*, etc.), ou une conclusion (*en conclusion*, etc.). Ils peuvent introduire d'autres éléments dans l'énoncé, comme un exemple (*par exemple*), une comparaison (*par rapport à*), etc.

Notre deuxième grand groupe concerne les marqueurs à fonction argumentative. Les marqueurs argumentatifs servent à construire le raisonnement et l'argumentation par des procédés d'explication, de justification, de rectification, de concession. Si les marqueurs de fonction métadiscursive participent à la structure globale du texte, les marqueurs de fonction argumentative contribuent à établir les liens internes du texte. Ces marqueurs renvoient dans la plupart des cas aux connecteurs argumentatifs. Garnier & Savage (2011) décrivent avec justesse les opérations fondamentales de l'argumentation. Quand les scripteurs rédigent leur

écrit, ils sont amenés à « expliquer, justifier, rectifier, concéder, souligner des contradictions, c'est-à-dire à accomplir un certain nombre d'étapes dans leur raisonnement » (*ibid.* : 106). Les marqueurs à fonction argumentative sont décrits dans des classifications proposées par Adam (2011), Riegel *et al.* (2009) et Garnier & Savage (2011) notamment. Nous distinguons des sous-groupes des marqueurs qui expriment la causalité, la concession, l'opposition, la condition et la finalité.

Notre classification des séquences lexicalisées à fonction discursive doit remplir les critères linguistiques et didactiques qui correspondent à des objectifs à la fois d'ordre linguistique et didactique. D'une part, sur le plan linguistique, les séquences polylexicales sont regroupées selon leur fonction discursive. Nous définirons ci-dessous chaque groupe dans notre classification des séquences lexicalisées à fonction discursive, et préciserons différentes propriétés qui les distinguent. D'autre part, sur le plan didactique, ce classement doit être aussi clair que possible, afin de permettre au lecteur d'accéder rapidement à la signification des séquences polylexicales sans recourir à des explications linguistiques complexes. Comme nous l'avons mentionné dans le chapitre 2, les séquences polylexicales sont moins abordées dans l'enseignement que les marqueurs de discours monolexicaux. De plus, les apprenants font de nombreuses erreurs dans l'utilisation de ces éléments. Dans le chapitre 6, nous allons réfléchir aux moyens qui peuvent être mis en place pour l'enseignement et l'apprentissage de ces séquences polylexicales (*Cf.* Chapitre 6 Corpus de mémoires d'apprenants vietnamophones). Enfin, la typologie que nous allons mettre en place servira de base de données pour les activités didactiques, que nous décrirons au chapitre 6.

La Figure 4-1 contient l'exemple d'un extrait d'un article scientifique où toutes les séquences lexicalisées à fonction discursive sont mises en gras et soulignées. Nous pouvons constater que ces éléments linguistiques sont très fréquents dans les articles scientifiques.

Deux principes généraux guident la sociologie mathématique des processus : la mathématisation **d'une part**, et le refus du descriptivisme **d'autre part**. Prenant pour acquis que toutes les sciences sont régies par le même régime épistémologique (ex. Fararo, 1969b, 75), les sociologues mathématiciens déplorent le faible niveau de scientificité de la sociologie. Celui-ci est dû, nous dit-on, à l'utilisation insuffisante du langage mathématique, **en particulier** en théorie sociologique. En recourant aux outils de la mathématique, les sociologues mathématiciens suggèrent d'élever le niveau de scientificité de la sociologie. Le langage mathématique est **en effet** jugé supérieur au langage naturel (Simon, 1957, 90). Il est apprécié pour sa rigueur, sa précision et sa puissance (Coleman, 1964, vii; Simon, 1957, viii). Ces qualités permettent un meilleur contrôle de la portée des propositions théoriques ainsi qu'une étude systématique de toutes leurs conséquences. **De plus**, il est considéré que les sciences sociales, qui sont particulièrement complexes (Coleman, 1964, 1), peuvent **d'autant plus** bénéficier des vertus du langage mathématique (Simon, 1957, 89). *[Extrait du texte 802, sociologie]*

Figure 4-1 : Exemple de l'utilisation des séquences polylexicales à fonction discursive dans un article scientifique

Nous allons maintenant détailler les neuf sous-groupes de notre typologie des séquences lexicalisées à fonction discursive.

4.2. Typologie des MD

Dans cette partie, nous allons tout particulièrement insister sur les marqueurs métadiscursifs. Ces marqueurs sont à notre avis moins abordés dans les manuels de FLE que les argumentatifs, et ils sont pourtant très fréquents dans notre corpus. Nous allons montrer que ces marqueurs d'organisation textuelle sont assez spécifiques des écrits scientifiques en SHS. Nous avons recensé au total 171 MD.

4.2.1. Marqueurs d'enchaînement thématique

Nous regroupons dans cette catégorie tous les marqueurs qui servent à introduire une idée, une nouvelle thématique ou une suite d'éléments. Nous nous focaliserons sur les marqueurs d'addition, les marqueurs d'énumération, et les marqueurs de topicalisation (terme emprunté à Combettes (1998)), ce qui totalise 41 marqueurs.

a/ Marqueurs d'addition

Les marqueurs d'addition, comme leur nom l'indique, servent à introduire un nouveau contenu qui peut correspondre aux arguments, aux faits, aux idées, etc. En fonction du rapport entre la séquence introduite par les marqueurs et le contexte gauche, nous pouvons distinguer deux types de marqueurs d'addition : 1/ ceux qui relient les unités qui se trouvent au même niveau, comme *entre autres*, *ainsi que* etc. ; 2/ et ceux qui signalent un changement de thématique comme *par ailleurs*, *de surcroît*, *en outre*, *au demeurant*, *en plus*, *de plus*.

Les éléments du premier sous-groupe, *entre autres*, *ainsi que*³³, ne portent que sur un syntagme nominal et se trouvent au milieu de la phrase [1][2][3]. Si *ainsi que* peut servir à introduire un dernier élément d'une série [1][2], *entre autres* suggère l'existence d'autres éléments, mais il souligne un seul élément d'un ensemble [3]. Pour les marqueurs du deuxième groupe, les informations qu'ils introduisent ne se trouvent pas au même plan que celles de leur contexte gauche [4][5][6]. Dans [4], le locuteur compare les résultats de son étude avec ceux d'autres recherches. La proposition où s'intègre *en outre* complète la séquence du contexte gauche en affirmant ce qui différencie son étude par rapport aux autres. Dans l'exemple [5], la proposition dans le contexte gauche de *par ailleurs* sert à introduire une précision sur les tâches de l'étude, tandis que le contexte droit apporte un jugement de l'auteur. Dans [6], le syntagme prépositionnel introduit par *de surcroît* sert à compléter et à renforcer ce qui a été dit dans le contexte gauche.

- [1] Le respect de l'ordre des épisodes **ainsi que** leur suivi continu ne sont pas primordiaux, donnant lieu à des consommations potentiellement moins assidues, plus opportunistes et davantage appropriées au visionnage à plusieurs. [infcom-art-21-body]
- [2] Le premier ensemble la région urbaine du Nord correspond grossièrement au delta du Fleuve Rouge, où la densité des centres urbains est élevée ; un triangle urbanisé, organisé autour de Hanoi, concernant les provinces de Yen Bai, Thai Nguyen, Vinh Phuc, Phu Tho, au nord, et celles de Hoa Binh et Thanh Hoa, au sud/sud-ouest, **ainsi que** le littoral de la province de Quang Ninh. [geo-art-7-body]
- [3] Il apparaît actuellement nécessaire de décrire ces troubles et les dissociations éventuelles rencontrées dans les profils pragmatiques, afin d'envisager, **entre autres**, la construction de tels outils. [lin-art-105-introduction]

³³ Il est à préciser que nous abordons uniquement le cas où *ainsi que* est utilisé pour ajouter une nouvelle information. Toutes les autres combinaisons de *ainsi que* sont exclues, par exemple quand *ainsi* sert à introduire une conséquence, il est suivi d'une proposition précédée par *que*, comme dans l'exemple suivant : C'est **ainsi que** des pratiques partagées et en apparence identiques (lever le doigt par exemple) supposent, pour être mises en œuvre de manière pertinente, une « règles de l'échange qui ne peut se construire que dans la pratique et la fréquentation régulière de celles-ci (un nouvel élève mettra ainsi plusieurs jours pour » « implicites de la règle » le doigt ». [sed-art-8-body]

- [4] Les résultats de notre étude sont non seulement conformes à ceux d'autres travaux récents (Atwell, Conners et Merrill, 2003 ; Vinter et Detable, 2003), mais ils montrent **en outre** une stabilité de l'effet de l'apprentissage implicite différé en fonction du degré de RM, avec des diminutions du pourcentage de respect du sRP d'environ 52 % chez les sujets avec RM léger et 63 % chez ceux présentant un retard plus important. [psy-art-53-body]
- [5] Précisons, en dernier lieu, que cette recherche expérimentale établit simplement des constats sur l'effet de transfert sans préjuger de la nature des processus sous-jacents ; mécanisme qui **par ailleurs** se révèle très complexe à interpréter et qui donne lieu à de nombreuses théories controversées. [sed-art-52-body]
- [6] Peirce, de son côté, et, comme nous le verrons plus loin, Wittgenstein, de l'autre, nous invitent à nous affranchir de cette maladie qui consiste à croire qu'on peut non seulement tout dire d'un objet, mais **de surcroît** dans un langage parfaitement précis et clair pour tous. [ant-art-19-body]

b/ Marqueurs de structuration

Nous regroupons les marqueurs *d'une part... d'autre part, d'un côté... d'un autre côté* parmi les marqueurs de structuration. Ces marqueurs servent à mettre en relief deux aspects ou deux propriétés d'un élément. Ils servent à repérer rapidement des idées développées par l'auteur et mettent en relation deux unités complémentaires [7][8][10] ou opposées [9].

- [7] Toutefois, il souligne que son fonctionnement correct supposerait une action commune, **d'une part** pour « garantir » aux États et aux entreprises que leurs partenaires ne recourront plus à des actions faussant la concurrence, et **d'autre part**, pour obtenir « la coopération sans réticence des travailleurs aux changements » (par des dispositifs permettant leur réadaptation) et compenser les handicaps de certaines régions (par un programme de développement en commun) [spo-art-10-body]
- [8] L'hypothèse que nous avons développée et qui vise à établir des relations entre **d'une part** des croyances pédagogiques, et **d'autre part**, les conditions sociales de production (positions, trajectoires, morphologie organisationnelle), n'est en rien contradictoire avec l'hypothèse selon laquelle chaque réseau aurait une culture pédagogique spécifique. [sed-art-28-body]
- [9] L'opérateur historique fait en effet face à un arbitrage lors de l'entrée d'un concurrent, puisque **d'un côté** il perd sur sa part de marché, mais **d'un autre côté** il récupère une partie de l'augmentation de la demande globale par la tarification de l'utilisation de l'infrastructure. [eco-art-279-introduction]
- [10] Si, **d'un côté**, la croyance méritocratique dans l'école, largement mobilisée par les enseignants les plus impliqués, se trouve renforcée, en particulier auprès des élèves qui suivent les ateliers CEP, **d'un autre côté**, cette présence contribue à changer leur vision des inégalités. [soc-art-641-body]

c/ Marqueurs d'énumération

Les marqueurs d'énumération servent à souligner la succession des éléments et des idées. Un segment textuel (c'est-à-dire une partie de la séquence, ou bien une ou des

propositions) peut être considéré comme l'un des termes d'une énumération. Les marqueurs d'énumération permettent de présenter non seulement le plan des parties textuelles, mais également l'enchaînement des idées. Les marqueurs de cette catégorie peuvent indiquer une introduction ou le premier élément d'une série (*au préalable, de prime abord, dans un premier temps, d'emblée, en premier lieu, etc.*). D'autres marquent les éléments intermédiaires (*en second lieu, par la suite, dans un second temps, etc.*). Ces marqueurs aident le lecteur à repérer facilement une suite d'idées développée par l'auteur. Certains marqueurs, par exemple *d'abord* ou *au final* peuvent également exprimer la valeur spatiale ou temporelle, valeurs qui sont exclues de notre étude [15][16].

- [11] Nous décrivons **au préalable** dans ses grandes lignes la stratégie mise en œuvre par l'analyseur pour résoudre les ambiguïtés du rattachement prépositionnel. [lin-art-259-body]
- [12] Nous reprenons succinctement **dans un premier temps** les principales spécificités de la modélisation ARFIMA et appliquons la technique d'estimation du maximum de vraisemblance exact aux séries mensuelles des variations logarithmiques des taux de change suivants : \$/ \$ canadien, \$/ Lire, \$/ Livre, \$/ Yen, \$/ Mark et \$/ Franc sur une période allant de Janvier 1974 à Juillet 1995. [eco-art-6-introduction]
- [13] En mobilisant ce corpus de sources, l'article cherche **d'abord** à montrer comment la réforme a été saisie par une élite professionnelle ; il s'attache *ensuite* à mettre au jour les répertoires d'action que celle-ci utilise pour mettre en œuvre la réforme ; il montre *enfin* les effets de la mise en place des PRES sur les fondements historiques du système académique français. [spo-art-4-body]
- [14] **D'abord**, nous devons nous interroger sur les rapports de parenté entre les Morend et les autres familles d'émeutiers : à ce propos, le lien le plus significatif renvoie aux « Bruchez 3 » de Villette (autour du sautier, François Bruchez) ; Pierre Bruchez, émeutier, était le cousin germain de Théodule Morend (leurs mères étaient sœurs) ; de plus, il était aussi son beau-frère, les deux ayant épousé deux sœurs Pellissier. [his-art-13-body]
- [15] [*]³⁴ Par exemple, des amis décident d'aller faire une virée en roller ou à vélo dans les rues de Paris : dans ce cas de figure, les participants ne s'affrontent pas, ils ne recherchent pas la victoire et si l'un d'entre eux décide de quitter le groupe dans leur épopée motrice, aucune sanction ne sera prise à son égard, car aucune sanction ni récompense ne sont prévues **au préalable** au sein d'un contrat ludomoteur. [sed-art-499-body]
- [16] [*] Dans le premier cas, le droit est **d'abord** associé à la société, il est un droit immanent, un law in action, susceptible d'être, comme le souligne Talcott Parsons lui-même un droit des individus, des collectivités privées contre l'arbitraire du pouvoir, de l'intervention gouvernementale (Parsons, 1962). [soc-art-35-body]

³⁴ [*] précède des exemples que nous n'allons pas prendre en compte dans notre corpus.

d/ Marqueurs de conclusion

Les marqueurs de conclusion sont utilisés à la suite d'une idée et servent à marquer le dernier élément d'une série (*en conclusion, pour finir, pour conclure*) [17][18]. Nous ne retenons pas *pour finir* ni *pour conclure* dans les exemples [19][20]. Dans ces exemples, ils ne fonctionnent pas comme un MD, mais plutôt comme une séquence libre dans l'association [Préposition + Verbe]. Nous distinguons les marqueurs de conclusion des marqueurs de reformulation servant à récapituler que nous verrons dans la partie suivante (*en somme, en résumé, en un mot*). Si les derniers servent à passer en revue toutes les idées dans leur contexte gauche, les marqueurs de conclusion introduisent un segment textuel ou un élément qui est considéré comme dernier d'une série d'idées.

- [17] **Pour finir**, nous chercherons à définir les causes sous- jacentes de la naissance de réseaux transfrontaliers dans les milieux nobles. [his-art-184-body]
- [18] Corrélativement, le voyeurisme trouve sa meilleure expression dans des contextes où l'obstacle d'un site hétérogène est éloigné du sujet S. **Pour conclure**, nous aimerions revenir sur le rôle d'interface que jouent les emplois perceptuels. [lin-art-136-conclusion]
- [19] [*] Dans le cadre de cet article, nous nous contenterons d'illustrer le passage de la structure linguistique à un patron utilisable automatiquement sur quelques structures définitoires : nous examinerons successivement les patrons des énoncés définitoires contenant un verbe du type définir (§ 3.1), les patrons des énoncés définitoires de signification (§ 3.2.1) **pour finir** par les patrons des énoncés définitoires de classification de forme Na est un Nx X (§ 3.2.2). [lin-art-169-body]
- [20] [*] Trois remarques **pour conclure** : Les exemples tirés de l'enquête MGIS ne sont pas anodins ni négligeables par rapport au reste de l'enquête. Les classifications des immigrés qui ont été discutées plus haut constituent l'armature de l'ouvrage qui présente les résultats de l'enquête et la polygamie est la première de toutes les questions examinées en détail.

f/ Marqueurs de topicalisation

Les marqueurs de topicalisation, aussi appelés les marqueurs de thématisation par Porhiel (2005), introduisent la thématique ou le sujet. Nous relevons 12 marqueurs dans cette catégorie. Ils regroupent les locutions adverbiales (*à cet égard, en ce sens*), prépositionnelles (*en guise de, à propos de, au sujet de, en ce qui concerne, en matière de, en termes de, pour ce qui concerne, sur le plan de, à l'exception de*) et des structures phrastiques (*il est question de*) [26].

En raison de leur structure syntaxique, ils sont en général suivis d'un syntagme nominal qui correspond au topique de l'énoncé. La particularité de ces marqueurs est, d'après Porhiel (*ibid.*), qu'ils « introduisent le thème en avant du texte » (*ibid.* : 113). Pour une

activité didactique, le repérage de ce genre de marqueurs est très intéressant. Ils servent à marquer un changement de thématique (Cf. l'emploi de *sur le plan de* dans [22]). Ils permettent aux apprenants de relever rapidement les thèmes abordés [23]. Par exemple : *en guise de* est toujours accompagné par des noms qui indiquent une partie textuelle du texte [21], comme *introduction, résumé, conclusion*, etc.

Les autres marqueurs dont un des composants est un déterminant démonstratif, ont une fonction anaphorique, et permettent d'attirer l'attention du lecteur sur une idée ou un élément dans le contexte antérieur (*à cet égard, à ce sujet, en ce sens*) [24][25]. Ils introduisent un thème au niveau phrastique et au niveau textuel.

- [21] Il est utile, **en guise de** conclusion, de résumer les caractéristiques des données du journal de bord qui les distinguent comme étant d'une valeur et d'un intérêt particuliers pour les climatologues. [his-art-152-body]
- [22] **Sur le plan de** la forme, la langue française est une langue flexionnelle riche, notamment à l'écrit, caractérisée par un décalage entre la morphologie de l'oral et celle de l'écrit (Jaffré, 1998). [lin-art-225-body]
- [23] Tous courts les plus longs font 7 pages, ces textes sont extrêmement différents **en termes de** l'approche : certains présentent la situation pour chacun des six objectifs de Bologne (i.e. la mise en place de la structure en deux cycles, du système de crédits ECTS, du supplément au diplôme, etc.) après une introduction générale (ou pas) sur le contexte national ; certains ne traitent que d'un objectif (en général, l'objectif pivot, l'introduction de la structure Bachelor/Master) ; d'autres présentent le système national d'enseignement supérieur sans renvoyer du tout aux objectifs de Bologne. [spo-art-117-body]
- [24] Il n'est pas nécessaire de recourir **à cet égard** aux techniques de régression multiple, dont on sait la difficulté à traiter de variables aussi fortement corrélées que celles de l'origine sociale et du niveau de diplômes et ce d'autant plus qu'elles correspondent à des caractéristiques logiquement et chronologiquement ordonnées : l'origine sociale intervient en amont du niveau d'études atteint et l'influence culturelle du milieu d'origine précède inévitablement celle de l'environnement scolaire. [soc-art-628-body]
- [25] Force est de constater **à ce sujet** que la conception qu'en donne Pierre Bourdieu, celle de « relations objectives » « que l'on ne peut ni montrer ni toucher du doigt, mais qu'il faut conquérir, construire et valider par le travail scientifique » (Bourdieu, 1994, p 9) ne fait pas l'unanimité. [soc-art-525-conclusion]
- [26] Quand **il est question de** gouvernance et non plus de gouvernement, quand l'excellence se substitue à la qualité scientifique, quand la compétitivité, devenue une fin en soi, remplace la compétition internationale pour la priorité de la découverte, on en conclut que la vision de la recherche construite au niveau national est désormais guidée par des concepts empruntés à une logique managériale. [soc-art-377-body]

4.2.2. Marqueurs de reformulation

Ces marqueurs permettent au locuteur de revenir sur son énoncé et d'orienter l'interprétation du sens par son interlocuteur dans le but de créer une meilleure compréhension. La reformulation peut se manifester par la reprise ou la mise en relief de ce qui précède en jouant sur la synonymie [27] et en mettant en relation des séquences équivalentes comme : *c'est-à-dire*, *en d'autres termes*, *autrement dit*, etc. D'autres marqueurs de reformulation servent à résumer ou à récapituler [28] les séquences précédentes, par exemple : *en somme*, *en résumé*, *en un mot*, etc. Reformuler demande des compétences spécifiques nécessaires à tout scripteur d'écrits universitaires. C'est pourquoi dans notre travail de thèse, nous souhaitons proposer une modélisation syntaxique et sémantique des marqueurs de reformulation et mener des réflexions sur les méthodes d'enseignement de ce type de marqueur. Nous avons relevé 12 marqueurs de reformulation dans notre corpus, qui feront l'objet d'analyses détaillées dans le chapitre 5.

- [27] Sur un mode ludique, la fiction leur permet de se glisser dans la peau des personnages fictifs, « non pas pour leur ressembler, mais pour vivre ce qu'ils vivent », **c'est-à-dire** pour enquêter sur les « belles affaires ».[infcom-art-5-body]
- [28] **En somme**, les trois hypothèses mises en avant par Hennin pour expliquer l'hostilité des représentants à la comédie (tactique, politique et sociale) fonctionnent toutes trois plus ou moins, mais il oublie celle qui les unit toutes, celle qui les subsume et que seule sa fureur trahit, à savoir que la comédie était devenue avant tout le symbole de la présence française. [his-art-315-body]

4.2.3. Marqueurs d'exemplification

Ces marqueurs permettent à l'auteur d'introduire les éléments servant d'exemples pour ce qui a été dit dans le contexte gauche : *par exemple*, *tel(le) que*, *à titre d'exemple*, *à titre d'illustration*, *en particulier*. Le marqueur *tel(le) que* est suivi uniquement par un syntagme nominal quand il sert à introduire un exemple, et l'adjectif *tel* doit être accordé avec le syntagme nominal qui suit [33]. *Tel(le) que* peut également porter sur une proposition, mais il n'exprime pas une exemplification dans ce cas [35]. Les marqueurs *à l'exemple de* et *à titre d'illustration* sont les moins utilisés dans cette catégorie et ont respectivement une fréquence de 10 et 8 occurrences [31][32].

Le marqueur *en particulier* permet au scripteur de mettre l'accent sur un élément, une idée, un fait ou un argument parmi d'autres [34]. Dans cette catégorie, *par exemple* reste le marqueur le plus utilisé (1357 occurrences). Il fonctionne comme un MD générique de cette catégorie [29]. *Par exemple* peut introduire la séquence qui sert d'illustration [29], de

justification [30], tandis que *tel(le) que* sert principalement à insister à un élément d'un ensemble [33], et *en particulier* permet au locuteur d'insister sur un élément important [34].

- [29] Les coopérations précédentes peuvent cependant tout autant faire figure de repoussoir : c'est, **par exemple**, le cas en Lorraine où le fonctionnement du pôle européen associant Nancy et Metz, jugé peu satisfaisant par les exécutifs universitaires nancéens, les incite à constituer une structure ne réunissant que leurs établissements. [spo-art-4-body]
- [30] Une telle perspective n'est pas surprenante dans la tradition américaine où le rôle de la Cour suprême est depuis longtemps au centre des préoccupations ainsi qu'en témoigne **par exemple** une analyse ancienne de Robert Dahl encore discutée aujourd'hui (Dahl, 1957). [soc-art-35-body]
- [31] **A titre d'exemple**, l'analyse du rôle différencié des réseaux d'information et d'observation dans les domaines de la protection marine, de la forêt et de la pêche permet de rendre compte du rôle structurant de cette forme d'instrumentation fondée sur l'articulation entre des instruments de type incitatif et des instruments de type informatif (voir tableau 2), tout en montrant ses limites pour l'autonomisation de l'action publique européenne.
- [32] **À titre d'illustration**, je voudrais rapidement évoquer une modeste expérience : j'ai été associé, à divers titres, à trois expositions à partir d'une enquête dans le monde des plantations de canne à sucre du Nordeste du Brésil : une à Paris, à l'École normale supérieure, une autre au Musée national, à Rio de Janeiro, la troisième, enfin, au petit musée municipal de Rio Formoso (État de Pernambouc), sur le terrain où je poursuis mon enquête.
- [33] Étant donné que le déplacement dans l'espace est soumis à certaines conditions de possibilités **telle que** la mobilisation du capital social, qu'il soit économique, temporel, relationnel ou culturel, ces différences de mobilité s'expliquent dans un premier temps par l'inégale répartition des ressources sociales entre les étudiants en mobilité internationale. [geo-art-6-conclusion]
- [34] Nous concluons sur le lien entre niveau des rendements et spécificités de carrière ; **en particulier**, les carrières des fonctionnaires sont plus régulières et plus courtes que celles des salariés du secteur privé alors que les espérances de vie sont plus longues, ce qui influe sur les rendements des régimes de pensions. [eco-art-9-summary]
- [35] [*] Cela étant, nous pouvons supposer que si le poids de l'étiologie avait dû se faire ressentir, la distribution **telle qu'**elle est aurait pu permettre son apparition, puisque le facteur RM lui-même ne différencie pas les performances d'apprentissage implicite. [psy-art-53-body]

Afin de vérifier si cette catégorie de marqueurs est spécifique aux disciplines de SHS, nous avons comparé l'utilisation de *par exemple* dans notre corpus avec un corpus représentatif de la langue générale, un sous-ensemble du corpus Emolex³⁵ d'une taille d'environ 40 millions de mots, composé des romans et des articles journalistiques. Nous relevons un ratio beaucoup plus élevé de *par exemple* dans notre corpus SHS (4,75 pour 10 000 mots) contre 1,01 pour 10 000 mots dans le corpus Emolex (4019 occurrences). En ce qui

³⁵ EMOLEX *le lexique des émotions dans cinq langues européennes : sémantique, syntaxique et dimension discursive*, piloté par Iva Novakova du LIDILEM, Université Stendhal – Grenoble 3 et Peter Blumenthal de l'Université de Cologne, site de l'Emolex : <http://www.emolex.eu/>

concerne *en particulier*, ce marqueur a un ratio de 2,03 dans notre corpus SHS (564 occurrences sur 2 678 811 millions de mots) contre 0,39 dans le corpus Emolex (1568 occurrences sur 40 millions de mots). Nous constatons que ces marqueurs sont donc beaucoup plus fréquents dans les écrits scientifiques. Dans un deuxième temps, nous avons comparé le ratio de *en particulier* de notre corpus SHS avec le corpus Scientext libre de droit³⁶ d'une taille plus grande que le nôtre (environ 4 900 000 mots) composé de quatre disciplines de sciences humaines (linguistique, psychologie, sciences de l'éducation et traitement automatique des langues) et de disciplines en sciences expérimentales et appliquées (biologie, médecine, électronique, mécanique). *En particulier* connaît un ratio un peu plus élevé dans notre corpus SHS : 2,03 pour un million de mots contre 1,68 pour un million de mots (830 occurrences) dans le corpus général Scientext. Nous pouvons ainsi confirmer que les marqueurs d'exemplification représentent une spécificité de nos écrits scientifiques en SHS.

4.2.4. Marqueurs de comparaison

Les marqueurs de cette catégorie servent à établir une relation entre deux ou plusieurs sujets entre le comparant et le comparé, entre son contexte gauche et son contexte droit. Nous regroupons dans cette catégorie des locutions adverbiales et prépositionnelles et des structures phrastiques (au total 12 marqueurs). Ils peuvent désigner une simple comparaison (*par rapport à*, *en comparaison de*, *comparativement à*, *par comparaison avec*). Le marqueur *par rapport à* est le marqueur le plus fréquent de cette catégorie (626 occurrences) dans notre corpus SHS [36]. Ces marqueurs peuvent introduire une comparaison entre les éléments semblables (*à l'instar de*, *autant que*, *de même que*, *il en est de même*, *il en va de même*, *au même titre que*) [37][38][39] ou non (*à la différence de*).

- [36] Terminologies et modèles conceptuels se démarquent d'ailleurs par le degré de cohérence des connaissances qu'ils représentent : le terminologue pourra s'arrêter à une validation locale du terme **par rapport à** des termes voisins, alors qu'en ingénierie des connaissances, les concepts introduits dans un modèle doivent maintenir la cohérence globale de ce modèle. [lin-art-50-body]
- [37] Dans cette optique, les Tic, **au même titre que** la présence sur le terrain, seraient un outil permettant d'expliquer le politique, et particulièrement les décisions, dans toute sa complexité. [incom-art-106-body]
- [38] Une façon d'interpréter ces résultats est de remarquer que, dans le long terme, la frontière des prix des facteurs implique que si le taux d'intérêt réel, fixé sur le marché international des capitaux, et le coût réel du travail non qualifié, fixé par le Smic, ne bougent pas, **il en est de même** du coût réel du travail qualifié. [eco-art-658-body]

³⁶ Consulté à cette adresse : <http://aiakide.net/scientext17/>

- [39] Nous poussons cette hypothèse nihiliste plus loin encore en affirmant qu'**il en va de même** pour les propriétés syntaxiques de sous-catégorisation. [lin-art-75-body]

4.2.5. Marqueurs de causalité

Les marqueurs de causalité instaurent un lien de cause à effet. Cette étiquette, qui est plus large que celle des autres catégories, regroupe aussi bien les marqueurs de cause que les marqueurs de conséquence et de déduction.

a/ Marqueurs d'explication

Certains marqueurs introduisent la cause et servent à apporter une justification et un exemple, un fait, un événement, par exemple : *parce que, en raison de, du fait de, étant donné que, d'autant plus que, ce qui signifie*, etc. [40][41][42].

- [40] D'une part, contrairement au système anglais, les relations ortho phonologiques en français sont relativement transparentes et offrent la possibilité d'établir les correspondances écrit oral sur de larges unités telles que les syllabes, **d'autant plus que** celles-ci sont clairement identifiables. [psy-art-303-body]
- [41] Cette valeur permet, de façon plus fiable que le simple calcul de la fréquence du couple (recteur, prép), de mesurer la force d'association entre les deux éléments, **dans la mesure où** une productivité élevée est l'indice que le recteur fonctionne de manière régulière avec cette préposition, donnant lieu à des occurrences diversifiées. [lin-art-75-body]
- [42] La lecture des forums de discussion associés à chaque article est en libre accès intégral, **ce qui signifie** qu'elle est ouverte à tous les lecteurs qu'ils soient scientifiques ou non. [infcom-art-432-body]

b/ Marqueurs de conséquence

Ces marqueurs introduisent une suite logique entraînée par un fait (*par conséquent, en conséquence, de telle façon que, c'est pourquoi, ce qui explique*, etc.) [43][44][45].

- [43] Nous proposons donc un modèle de signal au sens où la firme qui dispose de l'information privée sur la qualité du produit prend l'action observable par tous les consommateurs d'adopter ou non une politique de contrôle de la qualité et donc de produire **en conséquence**, et ce avant que les consommateurs aient pris leur décision d'achat. [eco-art-555-body]
- [44] En tout état de cause, ces trois exemples (dormir, courir et pleuvoir) montrent que, même lorsque le procès inclut une durée, ce n'est pas forcément sur ce trait que porte l'exclamation, **ce qui explique** sans doute le fait que certains verbes ponctuels puissent admettre la forme exclamative, par exemple cracher et fondre. [lin-art-274-body]
- [45] Les journaux de bord, **par contraste**, rendent compte de tous les événements, qu'ils soient extrêmes ou banals, et constituent de ce fait un enregistrement exhaustif de la fréquence et de la gamme des vents et événements climatiques. [his-art-152-body]

4.2.6. Marqueurs de concession

Les marqueurs de concession signalent une opposition ou une restriction à l'action exprimée par le verbe principal, par exemple : *en tout cas*, *par contre*, *en revanche*, *tout de même* [46][47][48][49], etc. Avec ces types de marqueurs, le locuteur peut revenir à la première formulation, soit pour la valider partiellement, soit pour l'invalidier. L'idée ou la conclusion dans la deuxième séquence peut apparaître comme étant en contradiction avec la première séquence.

- [46] On conclura des développements précédents que dans les situations de soumission forcée, la confrontation à une source de fausse attribution n'a pas, durablement **en tout cas**, immunisé les sujets contre la dissonance cognitive, sinon ces derniers n'auraient pas eu besoin de recourir, dans les recherches qui viennent d'être évoquées, à la voie de la rationalisation en acte ou à celle de la trivialisation. [psy-art-233-introduction]
- [47] Il est question, pour la Pologne, d'internationalisation, de modernisation (*modernizacja*), de développement de projets de construction (*investor*, investisseur ; planu, plan), de privatisation (*developer*, entrepreneur privé). **Par contre**, les grands ensembles est-allemands, à propos desquels la démolition et la déconstruction des cités sont mentionnées, sont qualifiés par leur besoin d'une revalorisation. [geo-art-192-body]
- [48] Mais vouloir mener une politique globale à l'échelle internationale en faveur du stockage du carbone, comme le préconise le Programme des Nations unies pour l'environnement (Pnue), participe **en revanche** d'une illusion assez persistante dans les instances internationales : celle d'une politique sectorielle répondant de façon universelle à un problème donné.[geo-art-23-conclusion]
- [49] C'est dans cette perspective que peut aussi se comprendre l'émergence, sur le terrain même des colonies, de nouveaux lieux de capitalisation, d'évaluation, de validation et de circulation des savoirs coloniaux. Les chambres d'agriculture implantées dans les colonies depuis les années 1760 ne bénéficient que d'un très médiocre prestige, mais s'attachent **tout de même** à collecter et à évaluer localement certains mémoires touchant notamment à l'agronomie. [his-art-216-body]

4.2.7. Marqueurs d'opposition

Ces marqueurs mettent en parallèle deux faits indépendants l'un de l'autre pour les opposer, par exemple : *au contraire de*, *par contraste* [50][51][52], etc. À la différence des marqueurs de concession, les marqueurs d'opposition expriment une opposition ayant trait à deux faits situés sur le même plan (*alors que*) [53].

- [50] Certains auteurs ont comparé l'écriture d'un article scientifique à un palimpseste (Artières, 2006) dont les premières versions sont écrasées par les suivantes, notamment au cours du processus d'évaluation par les pairs, lorsque les auteurs renvoient après chaque intervention des rapporteurs, un nouveau manuscrit modifié dans son intégralité, le superposant à l'ancien. Le dispositif proposé par la revue ACP permet **au contraire de** conserver de manière pérenne toutes les versions de l'article

ou plus généralement les deux ou trois versions successives de l'article ainsi que l'ensemble des commentaires réponses qui ont permis de passer de l'une à l'autre.[infcom-art-432-body]

- [51] Les questions soulevées par l'identification biométrique nous alertent sur la perte de sens à laquelle aboutirait en fin de compte une gestion purement techno-étatique de l'identité, soumise au seul impératif de fluidité et d'efficacité. La première conséquence, à bien des égards trop prévisible, d'un tel processus, est celle d'un décrochage d'avec la société réelle (d'une déprise institutionnelle), qui **par contraste** avec l'image irénique et totalisante du système, demeure un espace ouvert, poreux et hétérogène. [soc-art-162-conclusion]
- [52] On pourrait voir, dans cette personnalisation des règles de la vie sociale ou didactique, une tentation autocratique, voire narcissique du maître, **par opposition** aux valeurs démocratiques, fondées sur le dialogue et l'égalité, censées présider à la mise en œuvre des règles de vie ou des projets pédagogiques. [sed-art-8-body]
- [53] **Alors que** les modèles mathématiques ne permettent qu'une représentation superficielle des mécanismes (Manzo, 2010), les modèles computationnels (en particulier les modèles orientés objet) offrent la possibilité de représenter des mécanismes plus complexes (ne pouvant pas être résolu mathématiquement) ainsi que d'implémenter l'idée de générativité au sein des programmes de simulation (Hummon et Fararo, 1995, 86). [soc-art-21-conclusion]

4.2.8. Marqueurs de condition

Les marqueurs de condition introduisent la ou les conditions pour que l'action puisse se produire (*à condition que*) [54]. Ils peuvent introduire une hypothèse ou signaler le début d'une proposition où l'auteur décrit une hypothèse (*à supposer que*) [55].

- [54] C'est la reproduction (Bourdieu et Passeron, 1971) : le système se reproduit sans modification majeure, alors que certains acteurs peuvent être confrontés à des situations non conformes aux caractéristiques dominantes du système, **à condition que** leur nombre soit suffisamment faible par rapport à celui de ceux qui se retrouvent dans des situations conformes. [soc-art-98-body]
- [55] Finalement, **à supposer qu'il** existe bien une différenciation des aptitudes avec l'âge, ce sur quoi nous aurons l'occasion de revenir au cours de la discussion, les écarts d'âge entre les groupes (de 13 à 21 mois, cf. tableau II) sont, dans l'absolu, bien trop faibles pour constituer un facteur explicatif crédible, surtout si l'on tient compte du statut des participants. [psy-art-296-body]

4.2.9. Marqueurs de finalité

Ces marqueurs de finalité signalent l'objectif, le but de l'action : *afin de, de façon à, de manière à, à cet effet, en vue de, dans le but de, dans le dessein de, en sorte que* [56][57][58].

- [56] Cette unité d'analyse permet tout d'abord de résoudre une partie des écueils méthodologiques et analytiques posés par l'analyse de l'action publique européenne,

- notamment **en vue de** sa comparaison avec les politiques publiques nationales, voire avec le traitement de certains enjeux au niveau international. [spo-art-53-body]
- [57] Mais au-delà de l'effet de mode, la question d'un rattrapage du retard permet de justifier des initiatives publiques à la fois nationales et communautaires en faveur de la modernisation des entreprises européennes et de plaider pour la nécessité d'un effort dans la recherche **afin de** conquérir les bases d'un avantage à venir. [soc-art-523-body]
- [58] Le concept de « technique du corps », tel que Marcel Mauss l'a défini dès 1934, constitue **à cet effet** un outil précieux, car il invite à embrasser les pratiques concrètes dont le visage est le récepteur aussi bien que les intentions sociales qui les sous-tendent. [his-art-598-body]

Nous avons présenté ci-dessus les différents groupes des marqueurs dans notre typologie. Cette typologie n'est pas exhaustive, mais elle regroupe les séquences lexicalisées à fonction discursive les plus fréquentes de nos écrits scientifiques. Il est indispensable de les intégrer dans le programme d'enseignement. Nous avons vu que les écrits scientifiques en SHS sont caractérisés par la présence importante des marques de guidage (marqueurs de topicalisation, de structuration, etc.), et également les marqueurs métadiscursifs, c'est-à-dire les marqueurs d'exemplification, de comparaison, de reformulation.

4.3. Marqueurs de reformulation

Parmi les séquences lexicalisées à fonction discursive de notre typologie, nous avons fait le choix, pour des questions de faisabilité notamment, de ne travailler précisément que sur les marqueurs de reformulation. Nous aborderons toutefois pour les activités didactiques d'autres unités lexicales (*Cf.* Chapitre 6). Les modélisations des marqueurs de reformulation, ainsi que différentes activités d'enseignement et d'apprentissage mises en place, serviront principalement de modèle pour les adapter ultérieurement aux autres séquences lexicalisées à fonction discursive.

Dans cette partie, nous allons d'abord expliquer les raisons qui justifient notre choix pour les marqueurs de reformulation. Celles-ci sont relatives à leur fréquence, à leur rôle dans les productions écrites des apprenants. Nous allons passer en revue quelques typologies des marqueurs reformulatifs en français. Ensuite, nous présenterons la classe des marqueurs de reformulation retenue dans le corpus (Tableau 4.3).

4.3.1. Intérêt pour les marqueurs de reformulation

Notre intérêt pour les marqueurs de reformulation s'explique tout d'abord par leur fréquence très élevée dans les écrits scientifiques. Certes, les MD qui nous intéressent se retrouvent dans différents genres, mais nous nous intéressons spécifiquement à ceux qui sont surreprésentés dans les écrits scientifiques. En comparant la fréquence de quelques marqueurs de reformulation dans notre corpus et dans un sous-ensemble du corpus Emolex d'une taille d'environ 40 millions de mots, qui rassemble des romans et des articles journalistiques, nous constatons un ratio nettement plus élevé pour la plupart des MD sélectionnés dans notre corpus (Figure 4-2). Par exemple, *c'est-à-dire* connaît le ratio par 10 000 de mots de 2,53 pour notre corpus (704 occurrences sur 2,7 millions de mots) contre 0,35 dans le sous-corpus d'Emolex contenant des textes journalistiques et littéraires (1400 occurrences sur 40 millions de mots). On observe la même tendance avec *en d'autres termes* dont le ratio est de 0,19 pour 10 000 dans notre corpus contre 0,03 pour un million dans le corpus Emolex.

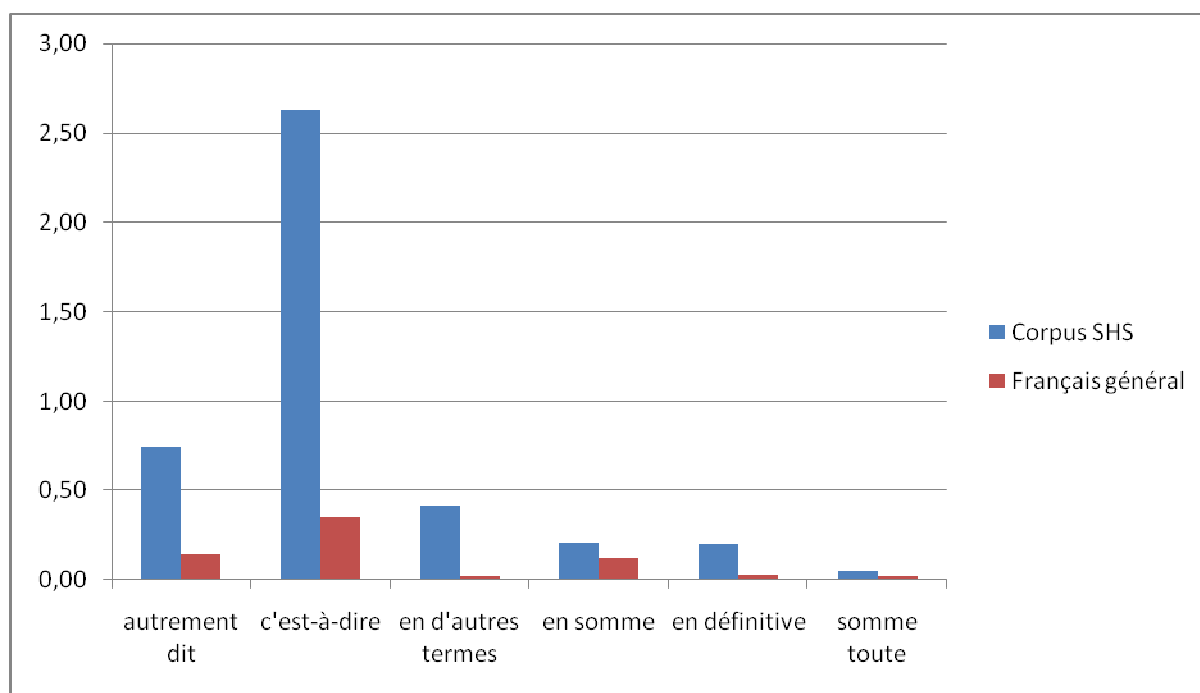


Figure 4-2 : Comparaison de l'utilisation des marqueurs de reformulation dans le corpus SHS et dans le corpus Emolex

En outre, notre choix est lié à des raisons d'ordre linguistique et didactique qui sont propres aux marqueurs de reformulation : leur rôle important de structuration du discours permet d'avancer qu'il faut les enseigner aux apprenants étrangers. Premièrement, Pétroff

(1984) observe que la reformulation est omniprésente dans le discours scientifique. Elle peut apparaître sous plusieurs formes à l'oral ou à l'écrit, qu'il s'agisse de la simple reformulation des idées, de phrases ou de la reformulation d'un ou des textes entiers pour produire des synthèses, des comptes rendus, des résumés, etc. La reformulation des mêmes idées scientifiques peut également s'exprimer sous la forme de formules ou d'images pour favoriser la transmission du savoir ou simplement pour favoriser la mémorisation. Il peut s'agir de la reformulation pédagogique dans la classe entre l'enseignant et les apprenants en vue d'une meilleure appropriation des connaissances. En bref, les formes de la reformulation varient et dépendent des objectifs de communication du locuteur. Précisons que nous n'aborderons pas ici la reformulation pour des finalités de vulgarisation, de changement de registre, de transcodage, car il s'agit, soit de cas spécifiques à l'oral, soit de cas largement étudiés dans le discours didactique (Bondhel, 1996 ; Kerbrat-Orecchioni, 2005).

Deuxièmement, sur le plan didactique, Mangiante & Parpette (2010) confirment également l'importance de la reformulation dans les compétences universitaires chez les étudiants dans le contexte du FOU. Selon eux, elle fait partie des trois macro-compétences transversales pour la production écrite, à savoir la compétence de re-produire (restituer le cours), re-formuler (altérer, argumenter) et re-lie (associer, approprier) des connaissances, des objets, des concepts. Mangiante & Parpette (*ibid.*) associent à chaque macro-compétence des tâches à effectuer par l'apprenant. En ce qui concerne la compétence à reformuler, il est demandé aux apprenants de développer la capacité de convaincre, d'argumenter, de commenter à partir de la prise de notes des cours et de démontrer.

C'est pourquoi la reformulation devient l'une des compétences indispensables à acquérir par le public visé – les étudiants, les jeunes doctorants, les apprentis chercheurs. En général, ils doivent rédiger des rapports, des mémoires, des thèses qui sanctionnent la validation d'un cursus ou d'un diplôme. Ces épreuves écrites sont suivies, dans la plupart des cas, d'autres épreuves orales, en l'occurrence des soutenances de rapport, de mémoire et de thèse. La reformulation passe par plusieurs formes, qu'il s'agisse de la communication des connaissances entre pairs, entre des chercheurs de même domaine, ou de la vulgarisation scientifique pour ceux qui ne sont pas du même domaine. Par exemple, les mémoires et les thèses s'adressent en premier lieu aux directeurs de recherche et aux membres du jury qui eux-mêmes appartiennent au domaine de spécialité ; les publications dans des revues spécifiques sont destinées essentiellement aux personnes du même domaine, etc. Il est

demandé aux étudiants d'apprendre à communiquer leurs travaux à un public homogène (à l'oral et à l'écrit) et à adapter leur discours au public auquel ils souhaitent s'adresser.

En outre, apprendre à reformuler facilite le processus d'appropriation des connaissances, des concepts et l'élaboration des communications. Cela demande aux étudiants de mettre en rapport ce qu'ils ont acquis auparavant et de réutiliser ces éléments ultérieurement dans d'autres contextes de travail. Effectivement, la reformulation établit le lien entre l'élément ou l'énoncé à reformuler et l'élément ou l'énoncé reformulatif. Pour que la reformulation soit efficace, les termes utilisés dans l'énoncé reformulatif doivent être précédemment acquis ou faire partie du savoir des lecteurs. Apprendre à reformuler profite aux étudiants sur deux plans : sur le plan textuel et sur le plan communicationnel. Sur le plan textuel, l'utilisation des marqueurs de reformulation est susceptible d'aider les étudiants à mieux structurer leur discours, leurs données dans le but de produire un discours cohérent. En outre, l'emploi des marqueurs de reformulation en particulier, et des séquences lexicalisées à fonction discursive en général, permet aux apprenants d'éviter la simple juxtaposition d'éléments, d'idées et de construire un discours cohérent. Au plan communicationnel, la reformulation manifeste le souci d'explicitier le discours de la part du locuteur. La reformulation permet d'établir donc l'interaction entre le locuteur et son interlocuteur.

Charolles & Coltier (1986) ont réalisé des tests auprès d'élèves de 5^e afin de comprendre l'utilisation des marqueurs de reformulation chez les apprenants et d'analyser le rôle des marqueurs de reformulation paraphrastique dans les processus rédactionnels. D'après eux, la production de marqueurs de reformulation est l'indice d'un « comportement coopératif » (*ibid.* : 58) et manifeste une volonté d'adaptation aux lecteurs de la part du scripteur. Les marqueurs de la reformulation paraphrastique permettent au locuteur de résoudre les problèmes de compréhension et également d'interprétation chez l'interlocuteur. Leurs conclusions confirment les fonctions décrites par Güllich & Kotschi (1983). Selon ces auteurs, la reformulation constitue un des moyens permettant de résoudre un certain nombre de problèmes communicatifs. Précisons toutefois que les genres ont des visées communicatives différentes selon les publics. Le genre scientifique ne remplit pas les mêmes rôles que les écrits demandés à des élèves de 5^e.

Par ailleurs, Charolles & Coltier (1986) et Thoiron & Béjoint (1991) soulignent l'importance de transmettre au lecteur la quantité d'informations strictement nécessaires pour qu'il arrive à comprendre et à tirer parti du texte. Le choix des informations dans la reformulation dépend fortement du contexte. L'usage des marqueurs de reformulation

demande avant tout au locuteur un certain nombre d'efforts liés au travail de composition textuelle, à savoir le choix d'un marqueur approprié dans le contexte, et l'effort lié à l'insertion de la séquence de reformulation dans le discours.

Plusieurs recherches insistent sur l'importance de la reformulation dans les textes scientifiques (Pétroff, 1984 ; Thoiron & Béjoint, 1991). Selon Pétroff (*op.cit.*), la reformulation consiste en « un acte essentiel du discours scientifique et technique » qui « reprend une information dans sa totalité et partiellement pour l'adapter soit à une situation nouvelle soit à d'autres destinataires » (*ibid.* : 54). Le discours scientifique est réalisé dans l'objectif de « réorganiser, contester, préciser, un réseau de connaissances antérieures pour un type de destinataires donné » (*ibid.* : 55), ce qui explique une présence abondante de MD. Ceux-ci permettent au locuteur d'établir la relation entre les idées et de construire la cohérence discursive entre les liens. L'explication des termes et des idées constitue une opération discursive de base dans le processus explicatif.

Après avoir passé en revue des rôles de la reformulation vis-à-vis de la production écrite des apprenants, il est maintenant important de donner une définition de ce phénomène linguistique. Les travaux fondateurs sur la reformulation de Gülich & Kotschi (1983), Roulet (1987), Auchlin (1981) s'inscrivent dans le cadre d'une linguistique textuelle orientée vers l'étude de l'interaction sociale. La fonction interactive est considérée, d'après les auteurs, comme un des critères importants dans la définition de la reformulation.

Cuenca & Bach (2007) considèrent la reformulation comme la co-élaboration du locuteur pour exprimer le contenu d'une partie du discours d'une manière différente. Il s'agit de reprises « aux fins d'explicitation, d'un concept C par une formulation Y différente de la formulation sous forme d'un terme X – et à leur utilisation dans les textes scientifiques » (Thoiron & Béjoint, *op.cit.* : 101). Dans beaucoup de travaux, la reformulation est décrite par le rapport de synonymie entre les expressions présentées comme équivalentes (Rossari, 1997 ; Riegel *et al.*, 2009). La définition de Charolles & Coltier (1986) correspond le mieux à notre perspective d'aide à l'écriture. Pour eux, l'analyse de la reformulation est particulièrement « intéressante pour l'approche des processus rédactionnels, car ces opérations sont révélatrices du travail de mise en texte effectué par le scripteur » (*ibid.* : 51). D'après leur définition, la reformulation apparaît :

dès qu'un locuteur/scripteur présente une expression comme explicitant (dans le contexte) la signification d'une autre : que cette explicitation (qui peut être une expansion ou une condensation) soit justifiée par référence au lexique (reformulations

proches de la dénomination) ou à un lieu commun (topos) comme c'est le cas dans les reformulations proches de la consécution... ou de la correction... (Charolles & Coltier, 1986 : 57).

Dans notre recherche, il s'agit de la relation entre le chercheur et ses pairs en vue de la communication écrite et orale de ses travaux. Avant tout, nous considérons la reformulation comme une stratégie de communication. Le scripteur recourt à la reformulation afin d'adapter son discours à un public déterminé en vue d'une meilleure compréhension entre les interlocuteurs et d'une acquisition du savoir transmis. Par le processus de reformulation, le locuteur est susceptible de revenir à l'énoncé formulé antérieurement dans le but de clarifier, compléter, interpréter et expliquer. La reformulation peut se réaliser par un lien d'équivalence entre des formes nouvelles et des formes connues de l'interlocuteur. L'énoncé reformulatif permet également d'introduire le point de vue du locuteur/scripteur. Nous allons analyser en détail les propriétés des marqueurs de reformulation dans le chapitre 5 sur la modélisation de ces éléments lexicaux. La liste des marqueurs de reformulation qui est retenue dans notre corpus se trouve dans le Tableau 4.3.

A l'issue de cette présentation des différentes études linguistiques sur la reformulation, nous souhaitons insister sur le rôle important de l'enseignement de la reformulation en général, et celui des marqueurs de reformulation en particulier pour l'aide à la rédaction scientifique. Cet intérêt est justifié par les études sur l'utilisation des marqueurs reformulatifs dans les productions écrites des étudiants (Boch, 2013 ; Rossari, 1992). Rossari (*ibid.*) a constaté une absence de connecteurs reformulatifs dans les rédactions d'étudiants. Boch (*op.cit.*) a effectué une analyse contrastive des pratiques (évocation, reformulation, citation) des experts et des étudiants en deuxième année de sciences du langage. Précisons ainsi que la notion de la reformulation dans ses recherches est un peu différente de la nôtre, dans la mesure où elle s'intéresse spécifiquement à la reformulation du discours d'autrui. Selon Boch (*ibid.*), la reformulation « permet au scripteur d'intégrer la parole de l'autre dans ses propos qu'il prend lui-même en charge » (*ibid.* : 558). Elle a relevé deux tendances contrastées : une forte présence de la reformulation chez les experts, tandis que dans les écrits des étudiants, la reformulation apparaît très peu pour laisser la place à la citation. Ces résultats sont-ils dus à un manque de connaissances sur le phénomène de la reformulation ? Ou sont-ils liés à la peur de reformuler le discours d'autrui ? Nous souhaitons donc avec ce travail sensibiliser les apprenants à ce phénomène linguistique, ainsi que mettre à leur disposition des moyens lexicaux pour faciliter leur tâche.

Nous concluons cette partie qui insiste tout particulièrement sur l'importance de l'enseignement/apprentissage des techniques de reformulation. Il existe trois grands types de reformulation dans les écrits scientifiques : reformuler son propre discours ; reformuler le discours d'autrui (un propos, un résultat, une hypothèse, etc.) ; reformuler pour communiquer les connaissances entre pairs. Dans notre travail, l'analyse de la reformulation désigne en général l'auto-reformulation dans le sens où le locuteur revient sur son énoncé pour le clarifier par d'autres moyens lexicaux. La reformulation peut concerner également la reformulation du discours d'autrui. Dans ce cas, il s'agit de la reformulation des travaux d'autres auteurs, par exemple, comme nous le verrons au chapitre suivant.

4.3.2. Différents classements dans d'autres recherches

Avant de décrire en détail différentes typologies qui inspirent notre travail, nous ferons un petit bilan des moyens lexicaux ou marques scripturales qui permettent d'introduire une reformulation. Comme l'a décrit Adam (1989), les opérations de reformulation peuvent être réalisées de différentes manières par la forme syntaxique ou par des marqueurs plus ou moins spécialisés. La reformulation est réalisée par des marqueurs qui se trouvent selon Adam sur un continuum allant des moins marqués aux marqueurs plus spécialisés. Les marqueurs les moins spécialisés consistent en une simple apposition manifestée par la ponctuation (parenthèses, couples de virgules, deux-points, etc.) [59][60].

- [59] Le point commun des figures choisies est d'être basées sur des références plurielles : il s'agit ici plus particulièrement de SN défini pluriel et de noms collectifs (Ncolls désormais) [ling--87-art-body]
- [60] Il existe deux grands types d'orthographe : celles que l'on appelle « transparentes », c'est-à-dire essentiellement régies par des correspondances graphèmes/phonèmes régulières, stables et univoques (finnois, italien...) et celles dites « opaques » ou « profondes » qui, tout en observant aussi le principe des correspondances phonographémiques, reflètent également les composantes morphosyntaxiques de la langue. [psy--109-art-body]

Boch (2013) a relevé différents moyens pour exprimer la reformulation du discours d'autrui dans les écrits scientifiques. Par exemple, la reformulation peut opérer avec ou sans marqueurs de reformulation explicites. Elle peut être introduite par les marqueurs introductives du discours rapporté (*selon...*, *d'après...*, *comme X prétend affirmer...* etc.) ou sans marqueur, car « le discours d'autrui est intégré dans le discours propre et n'a pas d'autonomie énonciative » (*ibid.* : 560).

Notre typologie concerne uniquement les moyens lexicaux qui servent à relier le terme reformulé et la séquence syntagmatique qui le reformule. Il s'agit de marques formelles telles que les éléments lexicaux (*autrement dit, à savoir, etc.*). À notre connaissance, il existe très peu de typologie des marqueurs de reformulation dans les écrits scientifiques. Nous analysons maintenant différentes typologies des marqueurs de reformulation qui serviront de base pour notre classement des marqueurs de reformulation. Plusieurs travaux soulignent l'existence de deux types d'opérations de reformulation : les marqueurs de reformulation paraphrastique (désormais MRP) (Gülich & Kotschi, 1983 ; Charolles & Coltier, 1986), et les marqueurs de reformulation non paraphrastique (Roulet, 1987 ; Rossari, 1990,1997).

4.3.2.1. Gülich & Kotschi (1983)

Gülich & Kotschi (1983) sont parmi les premiers à avoir travaillé sur les marqueurs de reformulation paraphrastiques (désormais MRP). D'après eux, ces marqueurs sont utilisés par les locuteurs pour marquer une reformulation. Ils représentent le moyen le plus important et le plus explicite pour marquer la relation de paraphrasage. Gülich & Kotschi (*op. cit.*) travaillent sur un corpus oral composé de divers enregistrements, en particulier la transcription d'un entretien téléphonique radiodiffusé et celle d'un cours de sémantique. Les MRP sont définis dans une relation de paraphrase, qui comporte trois constituants : l'énoncé-source, l'énoncé-doublon et l'élément qui indique une relation paraphrastique, c'est-à-dire le MRP.

Les MRP permettent d'établir une équivalence sémantique plus ou moins forte entre deux formulations par une relation de paraphrase qui résulte d'une « activité discursive de paraphrasage » (Kohler-Chesny, 1981, cité par Gülich & Kotschi, 1983 : 307). Ils n'entretiennent pas uniquement une relation de synonymie définie par « un type de relation interne au lexique d'une langue donnée » (Charolles & Coltier, 1986 : 55). Cette opération de reformulation permet au locuteur de rectifier ou clarifier sa première formulation avec le maintien d'un lien étroit entre deux formulations. C'est la relation paraphrastique entre deux énoncés qui détermine le choix des marqueurs de la reformulation paraphrastique et leur classement.

Comme la typologie de Gülich & Kotschi (*op.cit.*) est constituée d'un corpus oral, les MRP sont également très spécifiques de ce genre. Par exemple, les expressions complexes qui contiennent des verbes ou des substantifs qui renvoient au processus communicatif comme *expliquer, préciser, etc. je vais vous dire, pour préciser exactement ma pensée, etc.* ou bien

des morphèmes et locutions considérés comme des adverbes, des conjonctions, ou des interjections, etc., par exemple : *disons, d'accord, tu sais/vous savez*, etc. Nous ne trouvons évidemment pas dans notre genre des écrits scientifiques. De plus, il n'existe pas dans notre typologie des marqueurs qui servent à reformuler les énoncés d'un autre interlocuteur que le scripteur (*vous me dites que, ah ben, oui, mais écoutez*, etc.).

C'est plutôt le modèle d'analyse proposé par Gülich & Kotschi (*ibid.*) qui nous intéresse. En fait, ils proposent à la fois un traitement syntaxique et sémantique. Au plan syntaxique, Gülich & Kotschi s'intéressent à la position, aux relations structurelles entre l'énoncé-source, et l'énoncé-doublon. Ils se focalisent sur la relation sémantique entre les reformulations paraphrastiques pour en distinguer trois exprimant 1/ l'expansion (par des spécifications ou des explications) ; 2/ la réduction (par des synthèses ou dénomination) ; 3/ la variation quand les MRP indiquent la dissociation (par la récapitulation, la reconsidération) ou la correction (par le contenu ou la formulation). Cette typologie des relations sémantiques nous semble intéressante pour différencier un MD d'un autre.

4.3.2.2. Roulet (1987)

Roulet (1987) s'intéresse aussi aux opérations de reformulation qu'il appelle les « marqueurs de la reformulation non paraphrastique ». Par cette appellation, il ne cherche pas à opposer ces marqueurs aux marqueurs de la reformulation paraphrastique comme ceux qui sont étudiés par Gülich & Kotschi (*op.cit.*), mais à distinguer les uns des autres. D'après Roulet (*ibid.*), si les MRP permettent de lier deux constituants de même niveau hiérarchique et consistent en une simple paraphrase, les marqueurs de la reformulation non paraphrastique jouent un rôle important dans la réalisation de la « complétude interactive » de l'énonciateur. Autrement dit, l'énonciateur est amené à ajuster successivement son discours en fonction de la situation de communication et de ce qu'il présume du destinataire.

Roulet (1987) insiste tout particulièrement sur la propriété des reformulatifs à marquer le changement de perspective énonciative par rapport au discours antérieur, ce qui permet de les distinguer d'autres fonctions interactives de l'énoncé, entre autres les fonctions de type argumentatif. Le *connecteur reformulatif* ou le *reformulatif* (d'après l'appellation de Roulet) permet d'articuler deux mouvements discursifs dans une intervention. L'énonciateur intègre, suite au premier mouvement discursif, un nouveau mouvement qui subordonne rétroactivement un mouvement discursif antérieur. Celui-ci sert à l'interprétation rétroactive

du nouveau mouvement discursif. Le connecteur reformulatif signale le nouveau mouvement discursif. Le choix de différents reformulateurs entraîne différentes perspectives énonciatives.

En bref, les opérations de reformulation dans les travaux de Roulet (*ibid.*) visent davantage à « marquer un changement de perspective énonciative par rapport au discours antérieur qu'à reformuler (au sens étroit du terme) un constituant déterminé de celui-ci » (*ibid.* : 116). Selon lui, le changement de perspective énonciative se manifeste de trois manières différentes : 1/ pour présenter comme invalide la perspective énonciative adoptée antérieurement, 2/ pour préciser la nouvelle perspective adoptée par l'énonciateur et 3/ pour indiquer le type de changement de perspective opéré. En s'appuyant sur les différents aspects du changement de perspective énonciative, il distingue sept sous-groupes des connecteurs reformulateurs.

Nous n'intégrons pas dans notre classement les connecteurs reformulateurs des premier et deuxième sous-groupes de Roulet, car selon nous, *de toute façon, de toute manière, en tout cas* (1^{er} sous-groupe) peuvent avoir une valeur discursive et modale, mais leur valeur modale est plus fréquente. Il s'agit du même cas avec *en fait, en réalité, de fait* (2^e sous-groupe) qui peuvent se trouver à notre avis à la fois dans la classe des séquences lexicalisées à fonction discursive que dans celles à fonction modale.

Les cinq autres sous-groupes (dans le Tableau 4.1) indiquent le type de changement de perspective opéré et se distinguent par trois propriétés principales : la spécification de l'opération qui conduit à la nouvelle perspective énonciative, l'indication de la portée totale de l'opération et l'indication de la dimension temporelle de l'opération. Roulet (1987) résume les différentes combinaisons des propriétés comme dans le tableau 4.1.

Spécification de l'opération	Indication de la portée totale de l'opération		Indication de la dimension temporelle de l'opération
	-	+	
+	En somme Au fond	Somme toute Tout bien considéré Tout compte fait	-
	En fin de compte		+
-	Finalement En définitive	Après tout	

Tableau 4.1 : Différentes combinaisons de différentes propriétés (Roulet, 1987 : 120)

Le sous-groupe constitué de *au fond*, *en somme* (3^e sous-groupe) et celui de *en fin de compte* (4^e sous-groupe) précisent la nature et le terme de l'opération qui détermine le changement de perspective énonciative. Ces connecteurs n'indiquent pas la prise en compte de la totalité de l'opération. C'est la dimension temporelle de l'opération qui permet de distinguer l'un de l'autre. Alors que *en fin de compte* introduit l'acte principal présenté comme l'aboutissement d'un processus, l'utilisation de *au fond* résulte d'un examen approfondi et *en somme* permet de reformuler des éléments de même signe.

Le 5^e sous-groupe indique le changement de perspective énonciative qui résulte de la prise en compte de la totalité des éléments, *somme toute*, *tout bien considéré*, *tout compte fait*.

Finalemment, *en définitive* forment le sous-groupe des connecteurs reformulatifs qui introduisent la reformulation comme le dernier d'une succession de points de vue. Ils indiquent la dimension temporelle de l'opération qui détermine le changement de perspective énonciative.

Le marqueur *après tout* (7^e sous-groupe) peut indiquer à la fois une valeur temporelle et discursive. Celui-ci indique la portée totale de l'opération et sur la dimension chronologique, il résulte d'un examen de tous les éléments positifs et négatifs. En raison de la complexité de ces marqueurs, nous n'allons pas les retenir dans notre typologie.

Pour conclure, la prise en compte du changement de perspective énonciative qu'a adoptée Roulet (*ibid.*), représente une perspective intéressante, car il s'agit d'une propriété spécifique des marqueurs de reformulation non paraphrastiques. Nous allons nous inspirer de ce critère dans notre typologie. En revanche, il faudra envisager une autre approche pour ce qui concerne le traitement des valeurs sémantiques des MD, ainsi que le rapport entre les deux séquences établies par le marqueur. Ajoutons que la dimension temporelle de Roulet pour l'opération de la reformulation nous semble un peu abstraite et nous intéresse moins pour notre genre d'écrits.

4.3.2.3. *Rossari (1990, 1994)*

Rossari (1990,1994) s'intéresse à la fois aux MRP et aux marqueurs de reformulation non paraphrastiques. Pour ces derniers, elle a repris les marqueurs de la typologie de Roulet (*op.cit.*). Pourtant, son classement s'appuie sur le degré de prise de distance du locuteur. Dans le Tableau 4.2, les marqueurs non paraphrastiques sont classés de gauche à droite en allant d'une faible prise de distance du locuteur (*en somme*, *en bref*) à une prise de distance plus forte (*enfin*).

La perspective proposée par Rossari (1990,1994) s'avère intéressante, car elle propose de distinguer ces deux types d'opérations de reformulation selon le but poursuivi par le locuteur. D'une part, selon Rossari (1994), les MRP ont pour rôle principal de compléter, de clarifier et de rectifier ce qui a été formulé antérieurement tout en maintenant le lien étroit entre les deux énoncés ou les deux points de vue. Elle suggère de caractériser les marqueurs par leur propriété sémantique de prédiquer une identité entre les deux états de choses (*c'est-à-dire, en d'autres termes, autrement dit*). Par exemple, *c'est-à-dire* permet d'articuler deux formulations ou deux points de vue, même s'ils n'entretiennent aucune équivalence sémantique. Cela veut dire que l'état de choses évoqué dans la formulation introduite par le marqueur est considéré comme équivalent à celui évoqué dans la formulation antérieure.

D'autre part, les marqueurs de la reformulation non paraphrastique opèrent un changement de perspective énonciative du locuteur. Cela est dû à la « rétrointerprétation » de la formulation ou de l'énoncé introduit par le marqueur. Le changement de la perspective énonciative donne lieu à une prise de distance plus ou moins forte selon le choix des marqueurs.

Paraphrastique	Non paraphrastique			
	Récapitulation	Réexamen	Distanciation	Renonciation
C'est-à-dire En d'autres termes Autrement dit	En somme En un mot Bref	Tout bien considéré Tout compte fait Somme toute Après tout En fin de compte Finalement En définitive	En tout cas De toute façon De toute manière En fait De fait En réalité Au fond	Enfin

Tableau 4.2 : Des opérations de reformulation (Rossari, 1994 : 22)

Rossari (1990) distingue quatre sous-types d'opérations de reformulation non paraphrastique (Tableau 4.2). Comme nous l'avons précisé ci-dessus, nous allons exclure de

notre typologie des séquences lexicalisées à fonction discursive les opérations de distanciation (*en tout cas, de toute façon, de toute manière, en fait, de fait, en réalité*) et les opérations d'invalidation (*enfin*), car ces marqueurs ont à notre avis une forte dimension modale dans la mesure où ces séquences permettent d'exprimer le point de vue du locuteur sur le contenu propositionnel introduit. Contrairement à Rossari (*ibid.*), nous insistons aussi sur le traitement des caractéristiques syntaxiques des MD pour distinguer un marqueur d'un autre.

En résumé, les typologies des opérations de reformulation proposées par Gülich & Kotschi (1983), Roulet (1987) et Rossari (1994) sont pour nous des sources d'inspiration dans la conception de notre classement des marqueurs de reformulation. Comme nous l'avons évoqué, la perspective adoptée par Roulet (*ibid.*) mettant l'accent sur le changement de perspective énonciative nous semble très intéressante, d'autant plus qu'elle correspond bien à notre avis à la prise de distance du locuteur dans la typologie de Rossari (*ibid.*). En outre, à l'instar de Gülich & Kotschi (*ibid.*), nous envisageons d'adopter une analyse syntaxique et sémantique des marqueurs de reformulation dans ce travail. Nous nous trouvons dans un genre où nous ne retrouvons pas tous les types de reformulation du discours oral. Nous allons analyser uniquement les types de reformulation les plus pertinents dans notre contexte.

4.3.3. Notre classement

Nous allons présenter dans cette partie le classement des marqueurs de reformulation répartis en trois groupes en nous appuyant sur leur fonction dans le discours : les marqueurs servant à clarifier, à récapituler et à reconsidérer l'énoncé. Les marqueurs de reformulations les plus fréquents dans les écrits scientifiques figurent dans le Tableau 4.3 suivant :

Sous-groupes	Marqueurs	Fréquence	Répartition
Clarifier	C'est-à-dire	704	10
	Autrement dit	197	10
	À savoir	139	10
	En d'autres termes	119	10
Récapituler	En somme	57	9
	En résumé	21	9
	En un mot	14	7
Reconsidérer	En définitive	54	10

	Au total	49	9
	Somme toute	13	7
	Au fond	12	5
	En fin de compte	11	6

Tableau 4.3 : Liste des marqueurs de reformulation dans notre corpus SHS

Quelques marqueurs de reformulation qui apparaissent dans d'autres recherches (Roulet, 1987 ; Rossari, 1997) sont très peu utilisés dans notre corpus d'articles scientifiques, par exemple, *tout compte fait* (4 occurrences dans 4 disciplines), *après tout* (7 occurrences dans 3 disciplines), *au bout du compte* (7 occurrences dans 5 disciplines), *en bref* (7 occurrences dans 4 disciplines)). En nous référant à nos critères de transdisciplinarité et de répartition vus au chapitre précédent, ces marqueurs seront exclus de notre analyse.

Dans le Tableau 4.3, le premier groupe, qui correspond aux marqueurs servant à **clarifier** l'énoncé, comporte les marqueurs de la reformulation les plus utilisés (*c'est-à-dire, à savoir, autrement dit, en d'autres termes*). Comme nous l'avons analysé précédemment, la reformulation fait partie d'une des stratégies de communication du locuteur. Celui-ci manifeste sa préoccupation sur l'adéquation du terme choisi dans le contexte et souhaite revenir sur ce qui a été dit dans l'énoncé précédent dans l'objectif de l'éclairer, soit par l'utilisation de termes connus de son interlocuteur, soit pour synthétiser les idées évoquées.

Le deuxième groupe englobe des marqueurs servant à **récapituler** (*en somme, en résumé, en un mot*). Les marqueurs dans cette catégorie ne s'envisagent pas pour *résumer* ou « condenser (un texte, un discours) en peu de mots, en ne donnant que les informations principales » (d'après le TLFi), mais ils envisagent de « redire en résumant point par point, en s'en tenant à l'essentiel ; reprendre les points principaux » (d'après le TLFi).

Le troisième groupe recense les marqueurs servant à **reconsidérer** (*en définitive, au total, somme toute, au fond, en fin de compte*). Les marqueurs de cette catégorie sont moins fréquents que les marqueurs des autres catégories. Ils ont la particularité d'introduire le point de vue du locuteur, ce qui les distingue des autres groupes.

Conclusion

Dans cette partie, nous avons présenté notre typologie des séquences lexicalisées à fonction discursive. Cette typologie se base essentiellement sur les fonctions que jouent ces unités lexicales dans l'énoncé. Ces MD peuvent être regroupés en deux grandes catégories, les marqueurs métadiscursifs et les marqueurs argumentatifs. Si les premiers concernent la structure globale du texte, les derniers servent à exprimer le raisonnement logique. Le premier groupe joue plutôt le rôle d'organisation textuelle, le deuxième sert à véhiculer un contenu énonciatif.

Le classement fonctionnel des MD correspond à l'approche onomasiologique adoptée. Même si les MD dans notre typologie peuvent se trouver dans d'autres genres textuels, il est à souligner qu'ils sont surreprésentés dans notre corpus, et en particulier, les MD qui servent à exprimer l'exemplification, la reformulation, la comparaison, etc. qui sont caractéristiques des disciplines en SHS. En raison de leur fréquence dans ce genre d'écrits, il est nécessaire de sensibiliser les apprenants à leur rôle dans la structuration et dans l'organisation du discours. Dans cette recherche, nous insistons tout particulièrement sur les marqueurs de reformulation qui représentent, nous semble-t-il, un des moyens pour sensibiliser les apprenants aux techniques de reformulation. Ces marqueurs font partie selon nous des compétences à développer auprès d'un public universitaire. Dans le chapitre 5, nous allons procéder à la modélisation syntactico-sémantique des marqueurs de reformulation, modèle d'analyse qui pourra être adaptable à d'autres types de MD.

Chapitre 5.

Marqueurs de reformulation : Analyse syntactico-sémantique

5.1. Principes de la grille d'analyse

Même si nous ne pouvons dans ce travail de thèse que procéder à une analyse approfondie des marqueurs de reformulation, nous souhaitons proposer une grille d'analyse adaptable à d'autres MD. Il est donc important de bien définir la grille qui guidera cette recherche. Avant d'explicitier les paramètres de notre grille d'analyse des MD de reformulation, précisons tout d'abord les principes sous-jacents qui nous guident dans cette analyse.

5.1.1. Objectifs

Le choix d'un modèle linguistique s'avère compliqué en raison de la spécificité de notre objet d'étude – les MD polylexicaux. À la différence d'autres éléments lexicaux (comme les verbes, les noms, etc.), ce type de lexique nécessite d'être traité en tenant compte de son comportement discursif et sémantique (*Cf.* Chapitre 2 : Marqueurs discursifs). Dans ce travail, nous souhaitons tirer profit des descriptions syntaxiques et sémantiques des marqueurs de reformulation afin de mettre en place des suggestions didactiques pour l'enseignement de ces unités lexicales. La grille d'analyse choisie doit répondre à nos objectifs linguistiques et didactiques.

Sur le plan linguistique, cette grille d'analyse constituera le pilier principal pour guider les analyses linguistiques des MD dans ce travail. Ces descriptions linguistiques doivent être à la fois précises et claires, et mettre en évidence les spécificités d'emploi de chaque MD. Il est à noter que ces descriptions ne s'adressent pas directement aux enseignants de FLE, mais

serviront par la suite à concevoir des séquences d'enseignement/apprentissage des MD ou des descriptions simplifiées (Cf. Dicorpus Chapitre 6). Dans cette recherche, nous nous intéressons particulièrement à la transposition didactique qui sera destinée aux enseignants. Cette transposition nous paraît très importante puisqu'il existe encore peu de travaux qui allient les descriptions linguistiques et les réflexions didactiques à propos de l'enseignement/apprentissage du lexique spécifique des MD.

Sur le plan didactique, l'enseignant peut choisir d'attirer l'attention des apprenants sur telle ou telle propriété du MD et d'orienter l'apprentissage en fonction du niveau des apprenants, même si pour l'enseignement et l'apprentissage de ces unités lexicales, l'enseignant n'utilisera pas la même terminologie que dans une analyse linguistique. Nous reviendrons aux suggestions de transposition didactique des MD dans le chapitre 6, qui sera consacré aux réflexions didactiques.

Enfin, comme ce travail se situe dans le cadre de la linguistique de corpus, nous souhaitons profiter des avantages du TAL au service de l'enseignement/apprentissage des MD. La grille proposée doit aussi être adaptable aux modélisations en TAL, par exemple la création des grammaires de requête avec le système Scientext, ou des applications didactiques sur Dicorpus³⁷, un dictionnaire basé sur corpus que nous présenterons au chapitre 6. Dans ce chapitre, nous nous consacrerons essentiellement aux descriptions linguistiques.

5.1.2. Modèle de Paillard & Vu

Nous adoptons dans cette recherche le modèle d'analyse des MD proposé par Paillard & Vu (2012). Ces auteurs effectuent des analyses sémantiques de vingt-et-un MD en français. Il s'agit de MD assez fréquents dans les manuels de FLE, autant dans les communications écrites qu'orales. Ces MD sont regroupés en fonction de leur proximité formelle, par exemple *en fait*, *au fait*, *de fait*. Les auteurs procèdent, dans un premier temps, à une analyse de chaque MD qui s'effectue autour de six rubriques : la *formation lexicale*, la *sémantique lexicale*, la *portée*, la *position*, l'*orientation discursive* et la *propriété de détachement* ou *non-détachement*. Dans un deuxième temps, les MD de la même série sont mis en comparaison sur les plans sémantique et discursif afin de distinguer un MD d'un autre. Cette approche

³⁷ Il s'agit d'une base de données mise en place par Achille Falaise (ICAR, ENS de Lyon), cf. Tutin & Falaise (2013).

comparative nous semble intéressante en ce qu'elle permet de mettre en relief différentes valeurs et différents emplois des MD.

Dans ce travail, nous ne cherchons pas à décrire la *formation* de chaque MD comme Paillard & Vu (2012), ni à mener une étude étymologique comme Steuckardt & Niklas-Salminen (2005). Chaque MD sera analysé sur le plan syntaxique et sémantique à l'instar de Paillard & Vu quand ils cherchent à traiter les notions de *sémantique discursive* et de la *syntaxe discursive*. Selon eux, un MD accorde un statut discursif à la séquence sur laquelle le MD porte. La *sémantique discursive* du MD indique non seulement la sémantique du MD en tant que telle, mais encore le vouloir dire, l'intention du locuteur ainsi que l'état de choses du monde. Quant à la *syntaxe discursive*, Paillard & Vu (*ibid.*) s'intéressent à trois notions : la *portée*, la *position* et le critère de *détachement* ou *non-détachement*. Nous retenons uniquement dans cette recherche les notions de portée et position qui nous semblent plus intéressantes pour la modélisation des MD.

Les travaux de Paillard & Vu (*ibid.*) font partie des recherches qui allient l'analyse linguistique et l'aspect didactique pour l'enseignement et l'apprentissage des MD en FLE. Leur modèle linguistique est une source d'inspiration dans la conception des paramètres pertinents sur le plan didactique. Aussi, comme eux, nous ne nions pas l'influence sémantique de chaque constituant du MD sur le sens d'un MD en tant que tel et soulignons que la sémantique d'un MD n'est pas la somme des constituants du MD. Nous accordons aussi une grande importance aux aspects syntaxiques et discursifs des MD dans une perspective didactique.

5.2. Paramètres de la grille d'analyse

Afin de faciliter l'analyse des MD de reformulation effectuée dans cette partie et également de réadapter pour des analyses ultérieures d'autres MD, nous reprenons le modèle proposé par Paillard & Vu (2012).

La suite comportant les séquences lexicalisées à fonction discursive va être analysée selon le schéma suivant :

p MD q

Chaque composant du schéma peut être interprété de la manière suivante :

MD : marqueurs discursifs, dans notre travail, nous traitons uniquement des séquences lexicalisées à fonction discursive ; ce modèle d'analyse est tout à fait réadaptable pour les unités monolexicales ;

p : séquence qui précède la séquence lexicalisée à fonction discursive ou le **MD** concernant l'énoncé **p** ;

q : séquence qui suit la séquence lexicalisée à fonction discursive, la séquence **q** correspond à la portée du **MD**.

Dans cette analyse qui porte uniquement sur les marqueurs de reformulation, **p** correspond à la séquence reformulée, et **q** consiste en la séquence syntagmatique qui reformule **p**.

Nos analyses syntactico-sémantiques des marqueurs de reformulation vont s'articuler autour de quatre paramètres : 1/ **définition du MD** ; 2/ **portée syntaxique** ; 3/ **position syntaxique** ; 4/ **valeurs sémantiques**. Il nous semble important de définir tout d'abord le rôle de chaque paramètre avant de les appliquer pour analyser les marqueurs de reformulation qui figurent dans le Tableau 4-3 : Liste des marqueurs de reformulation dans notre corpus SHS. Nous finirons par aborder l'importance de l'association de ces paramètres dans la modélisation des marqueurs de reformulation en particulier, et des séquences lexicalisées à fonction discursive en général.

5.2.1. Définition

Comme le mentionne Rossari (1990, 1997), chaque MD a la propriété de prédiquer quelque chose. Cette rubrique est une entrée en matière et elle a pour objectif de donner une description du contenu sémantique des MD. Nous n'avons pas nous-même proposé nos propres définitions et avons choisi le TLFi *Trésor de la Langue française*³⁸ pour ses riches descriptions sur les valeurs sémantiques, étymologiques, historiques, etc.

D'éventuels synonymes des MD feront l'objet analyse contrastive dans les parties suivantes ; par exemple, *à savoir*, *autrement dit*, et *en d'autres termes* sont souvent suggérés comme synonymes de *c'est-à-dire*, même si ces marqueurs ne peuvent être remplacés par

³⁸ Le TLFi est développé par l'ATILF *Analyse et Traitement Informatique de la Langue Française*, de l'Université de Lorraine. Il est la version informatisée du TLF *Trésor de la Langue Française* – un dictionnaire de la langue française du XIX^e-XX^e siècle et consultable à l'adresse : <http://atilf.atilf.fr/>. Le TLFi offre aux utilisateurs trois niveaux de consultation: recherche simple, recherche assistée et recherche complexe.

c'est-à-dire que dans certains cas, ces trois marqueurs se caractérisant par des emplois différents. Nous y reviendrons lors de l'analyse de ces marqueurs.

5.2.2. Portée syntaxique

La notion de *portée* est une notion très féconde en linguistique. Elle se trouve au cœur de nombreuses recherches. La portée est considérée comme un élément essentiel dans les analyses sur la négation (Muller, 1991, 2005), pour les adverbes (Guimier, 1996), pour les cadres de discours (Charolles & Péry-Woodley, 2005 ; Sarda & Charolles, 2012) et pour les MD (Paillard & Vu, 2012). La définition de la *portée* varie selon les auteurs. D'un côté, la portée peut se référer à sa zone d'influence qui se limite soit à une part de l'expression sur laquelle l'opérateur exerce une influence (Allwood, 1997, cité par Danell, 1998 : 3), soit au domaine dans lequel cet opérateur peut agir (Muller, 1991). D'un autre côté, Huumo (1996) prend appui sur l'aspect sémantique de la portée et définit la portée des adverbiaux locatifs comme leur influence sémantique sur les constituants de l'énoncé³⁹ (notre traduction) (*ibid.* : 266). Ces différentes définitions de la portée révèlent donc deux propriétés, l'une syntaxique et l'autre, sémantique.

Étant donné le nombre important de travaux sur la portée, il nous semble important de dresser un petit bilan des travaux en la matière pour expliquer ensuite comment nous envisageons de traiter cette notion dans notre travail. Dans son étude, Guimier (1996) se base sur l'aspect syntaxique et sémantique des adverbes en *-ment* pour distinguer leur incidence de leur portée. Si l'incidence correspond à la caractéristique syntaxique de l'adverbe, la portée est relative à sa propriété sémantique. Selon lui, l'incidence d'un adverbe renvoie donc à la « nature du constituant de la phrase auquel il se rattache » (*ibid.* : 104). La portée d'un adverbe indique « l'élément à propos duquel l'adverbe dit préférentiellement quelque chose » (*ibid.* : 4). La notion de portée permet à Guimier de catégoriser et classer des adverbes exophrastiques. Ceux-ci sont définis comme « des constituants externes à la phrase [...] et représentent des traces de l'intervention du locuteur qui commente tout ou partie de son énoncé ou de l'acte qui le produit » (*ibid.* : 6). En s'appuyant sur les propriétés de l'acte de l'énonciation, Guimier distingue trois types de portée des adverbes exophrastiques, à savoir, la portée sur le dit, sur le dire et sur la visée du discours. Pour chaque type de portée, Guimier relève différents types d'adverbes. Par exemple, les adverbes évaluatifs et assertifs

³⁹ « Its semantic coverage other elements in the sentence » (Huumo, 1996 : 266)

(*heureusement, curieusement, probablement*, etc.) portent essentiellement sur le dit. Les conjonctifs, les adverbes métalinguistiques, paradigmatiques (*respectivement, finalement, autrement, textuellement*, etc.) portent sur le dire. Enfin, les adverbes de cadrage, et les adverbes illocutifs (*linguistiquement, plus généralement, récemment, réellement*, etc.) portent sur la visée du discours. À l'instar de Guimier, nous nous intéressons également à la propriété sémantique de la portée des MD qui servira par la suite à analyser leurs valeurs sémantiques. Cependant, le fonctionnement syntaxique de la portée nous semble important, car selon nous, cette propriété syntaxique est aussi discriminante et susceptible de différencier un MD d'un autre. L'analyse de la portée des MD doit aussi traiter cet aspect syntaxique.

Dans une perspective un peu différente de Guimier, Bonnot (2001) traite un autre aspect de la portée dont la *portée syntaxique*, procédé central dans l'étude de Bonnot (2001) sur les mots du discours. Selon elle, la portée est introduite par les mots du discours qui les commentent et spécifient « son statut par rapport à celui d'autres termes X », X »'... qui auraient été susceptibles d'occuper la même position dans la chaîne parlée » (*ibid.* : 10). Elle distingue trois types de portée des mots du discours qui correspondent à trois niveaux de fonctionnement, le niveau de la prédication, de la dénomination et de l'énonciation. Au niveau de la prédication, le mot du discours peut porter soit sur un terme instanciant une place d'une relation prédictive, soit sur la relation prédictive tout entière. Au niveau de la dénomination, le mot du discours porte sur une unité lexicale opposable à d'autres unités lexicales susceptibles de catégoriser un référent donné en le rattachant à telle ou telle notion. Au niveau de l'énonciation, le mot du discours peut porter soit sur le thème, soit sur l'énoncé entier, composante illocutoire comprise. Quand le mot du discours porte sur l'énoncé, il spécifie son statut par rapport à celui d'autres énoncés qui auraient pu être prononcés dans la même situation. L'examen de la portée des mots du discours (sa nature, les types de portée acceptables à chaque niveau) permet à Bonnot de délimiter la classe qui les forme et de définir les facteurs de variation régulière indépendants de leur sémantique. En s'appuyant sur différents niveaux de portée, Bonnot a proposé une typologie des mots du discours. Cette recherche révèle de nouveau l'intérêt de travailler sur la portée des séquences lexicalisées à fonction discursive dans l'objectif d'une modélisation sémantique. Nous rejoignons Bonnot (2001) sur l'importance de l'analyse de la portée des MD qui ne doit pas se limiter à la nature de celle-ci, mais doit également traiter son statut dans l'énoncé, c'est-à-dire sa position dans l'énoncé. Dans notre recherche, une telle analyse des aspects syntaxiques et sémantiques des MD permet de relever des régularités de fonctionnement de ces éléments linguistiques.

Le nombre important de travaux sur la *portée* prouve sûrement qu'il s'agit d'une notion très complexe. Comme la portée est souvent liée à la spatialité (*scope* en anglais), Charolles & Vigier (2005) se sont intéressés à la portée des adverbiaux en position préverbale dans la structuration du discours. Ils définissent la *portée cadrative* comme l'ensemble des phrases qui appartiennent toutes à un même cadre de discours (*ibid.* : 9). La *portée cadrative* révèle « la capacité qu'ont certains adverbiaux antéposés d'étendre leur influence au-delà de leur phrase d'accueil » (*ibid.* : 9). Charolles & Vigier ont montré que certains adverbes sont plus aptes que d'autres à posséder une portée cadrative, par exemple, les adverbiaux scéniques (de temps ou de lieu) ou les médiatifs (*selon X*, etc.). Dans notre travail, la portée cadrative est moins adaptée nous semble-t-il à notre objet de recherche, car nous nous intéressons à des aspects plus immédiats de l'utilisation des MD, telle que la nature syntaxique de la séquence introduite par les MD ainsi que leur fonction dans la structuration du discours. Nous resterons donc sur l'analyse de leur portée syntaxique et sur celle de leurs différentes valeurs sémantiques.

Après avoir passé en revue différentes recherches sur la *portée*, nous pouvons conclure que cette notion est pertinente pour notre travail. Nous retiendrons que la portée **q** des séquences lexicalisées à fonction discursive est un des constituants de l'énoncé. Cet élément dépend de l'utilisation des marqueurs correspondants. Ces marqueurs mettent en relation son contexte gauche et sa portée, c'est-à-dire entre **p** et **q**.

$$[\text{ }_p \text{ MD } [\text{ }_q$$

Par la *portée syntaxique* des séquences lexicalisées à fonction discursive, nous pouvons déterminer la nature de la séquence introduite par le MD. Nous supposons que certains marqueurs portent, de façon privilégiée, sur une phrase ou sur une proposition alors que d'autres porteront sur des constituants plus réduits. Afin de faciliter la modélisation des séquences lexicalisées à fonction discursive, nous nous inspirons de l'analyse de la portée de Paillard & Vu (2012) qui en distinguent deux types : la *portée globale* et la *portée locale*. Si le MD porte sur une proposition ou un énoncé, alors il s'agit d'une *portée globale* ; en revanche, si le MD porte sur un syntagme, alors il s'agira d'une *portée locale*. Dans les exemples suivants, le MD *autrement dit* a une portée globale, il porte sur une proposition [61], tandis qu'*à savoir* prend sous sa portée un syntagme nominal, il a une portée locale [62].

- [61] D'autre part, lorsque nous interrogeons les élèves qui adoptent un sport au sein de clubs sportifs, nous nous apercevons que leur motivation n'est pas toujours au rendez-

vous quand il s'agit de les pratiquer en EPS. **Autrement dit**, ces élèves peuvent être démobilisés par l'utilisation modifiée et simplifiée de leur sport. [sed-art-52-body]

- [62] Malgré toute la bonne volonté du monde, s'intéresser à la participation des populations locales, à moins de s'inscrire dans une longue durée, c'est envisager des situations de dépendance, c'est-à-dire de passer par des intermédiaires pour accéder à l'objet étudié, **à savoir** les populations. [geo-art-575-conclusion]

Nous retenons dans ce travail ces notions de portée locale et portée globale auxquelles nous allons attribuer les étiquettes suivantes : **P_G** (portée globale) et **P_L** (portée locale). En outre, nous supposons que le contexte gauche des MD est également un élément essentiel dans l'analyse des MD. Il est nécessaire de relever la nature syntaxique de la séquence **p** ainsi que le rapport entre celle-ci et la portée du MD, afin de mieux comprendre les valeurs sémantiques du MD. Pour faciliter la modélisation des MD en général et des séquences lexicalisées à fonction discursive en particulier, nous procédons à la schématisation de leur relation. Par exemple, dans la figure ci-dessous qui représente la portée d'un MD, nous pouvons interpréter que le MD peut porter à la fois sur un syntagme et une proposition et que son contexte gauche – la séquence **p** correspond soit à un syntagme soit à une proposition.



5.2.3. Position syntaxique

La plupart des séquences lexicalisées à fonction discursive retenues dans notre typologie sont des groupes adverbiaux, une catégorie qui a pour caractéristique une grande mobilité : ils peuvent occuper différentes positions dans la phrase. Ils peuvent être antéposés, insérés ou postposés. Ils peuvent porter sur un verbe, un adjectif ou un autre adverbe, etc. Pourtant, les adverbiaux ne sont pas utilisés d'une manière aléatoire. Plusieurs études ont montré que les différentes positions phrastiques des adverbiaux dépendent des choix stratégiques et des objectifs communicatifs de l'auteur afin de construire un discours cohérent (Charolles, 1997, 2003). Il nous semble donc important d'analyser la position des MD pour relever leurs règles d'utilisation. En termes discursifs, ce paramètre nous permet de comprendre l'intention communicative de l'auteur.

La position initiale correspond à la position en tête de phrase ou en zone préverbale et consiste en une position stratégique dans de nombreuses recherches (Charolles & Péry-Woodley, 2005 ; Ho-Dac, 2007). Selon Virtanen (2004), cette position initiale est dotée de trois fonctions importantes, celle de point d'ancrage, de point d'orientation pour le segment à venir ou d'indicateur d'informations cruciales (Virtanen, 2004 : 80). Charolles & Prévost (2003), Charolles & Péry-Woodley (2005) se sont intéressés au rapport entre la position des adverbiaux et leur portée dans la structuration du discours. Ils ont montré que les adverbiaux, s'ils sont en position initiale, n'occupent pas la même fonction organisatrice que quand ils se trouvent en fin de phrase. Les adverbiaux antéposés se caractérisent par une portée extraphrastique. D'après Péry-Woodley (2000), les circonstants (propositions circonstanciellles ou adverbiales) à l'initiale ont généralement une portée qui dépasse l'énoncé. Les groupes adverbiaux contribuent fortement à la cohésion du discours en signalant le lien thématique ou référentiel avec le discours précédent et en regroupant plusieurs propositions.

Contrairement à Charolles & Péry-Woodley (2005), Crompton (2006) montre que les adverbiaux en position finale ont le même rôle organisateur du discours que les adverbiaux antéposés et peuvent étendre leur portée au-delà de la phrase. Son travail porte sur trois types d'adverbiaux : les adverbiaux de lieu et les temporels (ou en anglais « *adverbials of location* » (*ibid.* : 256)) (propositions temporelles formées par des prépositions comme *in, ago, during* etc.), les propositions infinitives de but et les propositions de condition constituées par *if*.

Ces différentes recherches ont pour point commun de confirmer l'importance des MD dans la structuration du discours. Dans notre recherche, nous analysons leurs différentes positions. Plus précisément, nous allons étudier la position du MD en relation avec sa portée. Ceci permettra de relever des emplois spécifiques des MD.

En ce qui concerne les marqueurs de reformulation faisant l'objet des analyses détaillées dans la partie suivante, ils introduisent toujours la séquence reformulatrice qui correspond à leur portée. Ils se trouvent essentiellement en position initiale, quelle que soit la nature syntaxique de **q**. Au niveau phrastique, ils peuvent se trouver en tête de phrase ou en position préverbale ou postverbale. Ce n'est pas le cas d'autres MD qui peuvent se trouver aux différentes positions comme *toutefois, cependant, d'ailleurs*, etc. Par exemple, dans les exemples suivants, *d'ailleurs* peut occuper la position initiale [63], médiane [64] et finale [65].

- [63] **D’ailleurs**, plusieurs cas observés de vieilles personnes célibataires ou précocement veuves sont exemplaires d’une vieillesse autonome qui se nourrit d’un bon réseau social au sein duquel les voisins ont toute leur part. [soc-art-59-conclusion]
- [64] Deux modalités de téléchargement sont **d’ailleurs** à distinguer : l’une concerne les séries relativement anciennes, qu’elles soient déjà terminées ou à un stade avancé. [infcom-art-21-body]
- [65] Débats plutôt anglophones que francophones **d’ailleurs**. [sed-art-124-notes]

Nous pouvons représenter la position de ce marqueur d’une manière suivante :

- En position initiale : []_p **d’ailleurs** []_q
- En position médiane : []_p [Début **q d’ailleurs** Fin **q**]
- En position finale : []_p [**q d’ailleurs**]

Certes, la position des marqueurs de reformulation est peu variée par rapport aux autres MD, mais elle constitue un paramètre important pour appréhender le comportement syntaxique du MD dans l’énoncé. Dans l’objectif de proposer un modèle d’analyse adaptable aux MD en général, il nous semble important de traiter la *position syntaxique* dans la modélisation de ces éléments linguistiques. Une analyse sémantique des MD permettra par ailleurs de mettre en valeur leurs différentes acceptions et de distinguer un MD de l’autre.

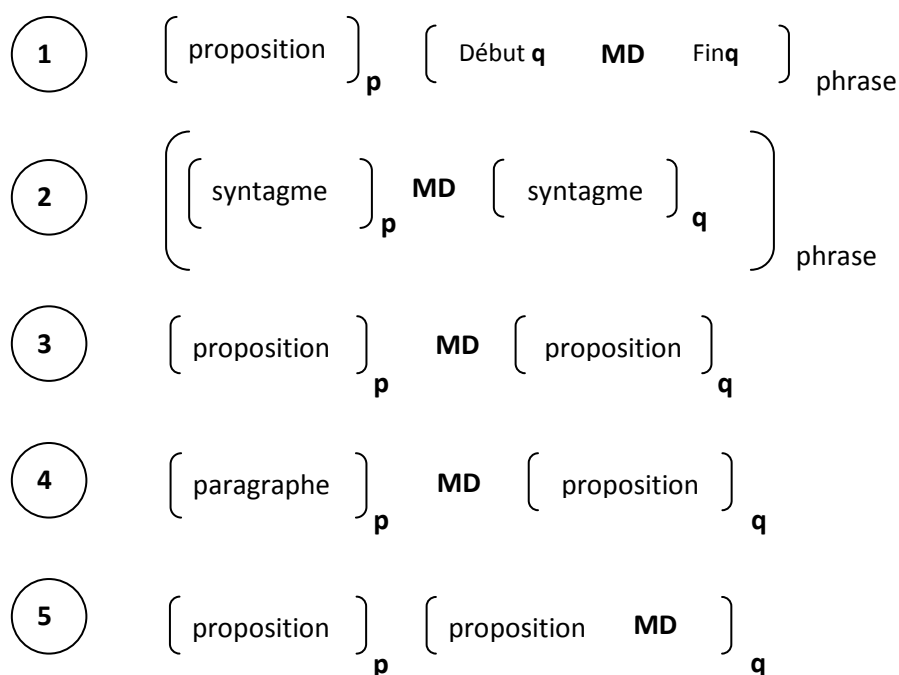
5.2.4. Valeurs sémantiques

Outre la description de la portée au plan structurel, l’examen précis de la nature sémantique du MD ainsi que du rapport avec le contexte gauche permet de relever ce qui différencie un MD d’un autre appartenant à une même classe syntactico-sémantique. Rossari & Jayez (1997) indiquent avec justesse que les MD ne peuvent que se différencier par leur noyau sémantique et que l’analyse de la portée permet de « faire ressortir des régularités quant aux configurations syntactico-sémantiques » (Rossari & Jayez, 1997 : 235). En effet, la portée d’un marqueur ne s’utilise pas d’une manière aléatoire. C’est pourquoi l’analyse du fonctionnement de la portée permet de délimiter les propriétés régulières des MD, ce qui est tout à fait indispensable dans leur modélisation. Prenons l’exemple de *c’est-à-dire*, *autrement dit*, *en d’autres termes* qui sont souvent utilisés comme synonymes les uns des autres. Nous ne nions pas que ces marqueurs puissent être interchangeables dans certains contextes, mais chacun contient ses propres contraintes de compatibilité. En outre, le changement de MD entraîne une variation sémantique dans le contexte de droite. L’analyse de la portée

sémantique nous sert à faire ressortir ces différentes nuances de sens et également à délimiter les conditions d'utilisation des MD. Elle est d'autant plus indispensable qu'elle permettra aux enseignants de mieux comprendre le fonctionnement de chaque MD.

L'analyse sémantique passe par une analyse empirique des propriétés sémantiques des MD, observées à partir des résultats de requêtes sur notre corpus informatisé de SHS. Nous travaillons également sur les corrélations entre le contexte gauche et le contexte droit du MD, c'est-à-dire entre **p** et **q** afin de révéler la valeur sémantique du MD dans chaque contexte. Nous supposons qu'il existe un fort lien entre **p** et **q** et que cette relation prédit quelques valeurs d'usage du MD.

Afin de mieux faire ressortir le fonctionnement syntaxique et sémantique des MD, nous introduisons à la fin de l'analyse linguistique de chaque MD un tableau qui récapitule ses traits syntaxiques et sémantiques essentiels. Nous pouvons représenter sous forme de schéma le rapport entre la position du MD et sa portée comme suivant :



Ce schéma représente les différentes positions du MD dans l'énoncé et par rapport à sa portée **q**. Nous pouvons interpréter ce schéma ainsi :

1. Le MD se trouve en position postverbale et la séquence **p** correspond à une proposition.

2. Le MD a une portée locale, il porte sur un syntagme et se trouve en position postverbale. La séquence **p** est un syntagme.

3. Le MD est en position initiale et a une portée globale, il porte sur une proposition et la séquence **p** est aussi de nature propositionnelle.

4. Le MD est en position initiale et a une portée globale, il porte sur une proposition et la séquence **p** peut se composer d'un ou même de plusieurs paragraphes.

5. Le MD est en position finale, sa portée et la séquence **p** sont des propositions.

Quant aux valeurs sémantiques, nous allons attribuer au MD des étiquettes correspondant à ses propriétés sémantiques qui sont notés en petites majuscules. Nous pouvons représenter ces valeurs sémantiques de la manière suivante :

$$\left(\begin{array}{c} \\ \end{array} \right)_{\mathbf{p}} \quad \mathbf{MD} | \text{VALEUR SÉMANTIQUE} | \quad \left(\begin{array}{c} \\ \end{array} \right)_{\mathbf{q}}$$

Pour résumer, la grille d'analyse que nous venons de décrire va nous guider dans l'analyse des marqueurs de reformulation. Elle sert à traiter les aspects syntaxiques et sémantiques de ces unités lexicales. Cette grille sera appliquée à l'analyse des séquences lexicalisées à fonction discursive pour des études ultérieures.

Analyse des MD de reformulation

Les descriptions linguistiques s'avèrent indispensables en apportant des explications sur des phénomènes linguistiques. Elles peuvent être adaptées plus tard pour les enseignants FLE et doivent être didactisées pour pouvoir enseigner dans la classe de langue.

Passons maintenant aux analyses syntactico-sémantique des marqueurs de reformulation qui se divisent en trois groupes comme nous l'avons précisé dans le chapitre précédent (Cf. Tableau 4-3 : Liste des marqueurs de reformulation dans notre corpus SHS). Nous allons appliquer les paramètres présentés ci-dessus pour l'analyse des MD. Dans chaque groupe, nous choisissons le marqueur le plus fréquent pour une analyse détaillée. Les autres marqueurs du même groupe seront mis en comparaison avec celui-ci afin de faire ressortir les spécificités de chaque marqueur. Nous effectuons essentiellement des tests de remplacement

afin de vérifier si les MD sont utilisés l'un à la place de l'autre et dans quels cas ils sont interchangeables.

5.3. Clarifier

Pour clarifier l'énoncé antérieur, le locuteur peut recourir à une autre formulation de même signification que celui-ci. Les marqueurs *c'est-à-dire*, *à savoir*, *autrement dit*, *en d'autres termes* permettent d'effectuer une reformulation adéquate du propos en établissant une certaine équivalence entre deux contenus. Ces marqueurs sont parfois mal utilisés par les apprenants étrangers, l'un à la place de l'autre en raison de leur sémantisme relativement proche (Cf. 6.1.2. sur l'analyse des erreurs les plus fréquentes dans l'utilisation des marqueurs de reformulation dans le corpus d'apprenants).

Dans cette classe, *c'est-à-dire* est le marqueur le plus fréquent, donc le cœur de notre comparaison. Les autres marqueurs *à savoir*, *autrement dit*, *en d'autres termes* seront comparés avec *c'est-à-dire*.

5.3.1. C'est-à-dire

C'est-à-dire reste le marqueur de reformulation le plus utilisé dans notre corpus, au total 704 occurrences réparties dans 10 disciplines. Force est de constater que ce marqueur fait l'objet de nombreuses études (Steuckardt & Niklas-Salminen, 2005 ; Vassiliadou, 2008 ; Khatchatourian, 2008). Ces descriptions très détaillées sur le plan énonciatif ou diachronique contribuent à mieux comprendre les valeurs sémantiques de ce marqueur. Il nous paraît nécessaire de dresser en premier lieu un petit bilan de différents travaux sur *c'est-à-dire* avant d'effectuer notre propre analyse linguistique.

Le modèle d'analyse des marqueurs de reformulation formés à partir du verbe *dire* que propose Khatchatourian (2008) nous inspire dans le traitement linguistique de ces éléments linguistiques. Dans sa recherche, elle traite essentiellement deux marqueurs, *c'est-à-dire* et *cela veut dire*. Khatchatourian propose de les analyser en s'appuyant sur les problèmes principaux de la communication concernant l'interprétation de ce qui est dit et la production de ce qui est à dire. Elle en distingue deux utilisations principales : quand la reformulation correspond à une définition exacte ou une interprétation du locuteur, et quand la reformulation est caractérisée par la subjectivité du locuteur. L'analyse des différentes stratégies de

communication nous paraît une approche intéressante pour relever les conditions d'emploi du marqueur *c'est-à-dire*, ainsi que pour comprendre les rapports entre les séquences établis par *c'est-à-dire*. Une analyse du sémantisme du verbe *dire* sert à mieux comprendre les valeurs d'emploi du marqueur.

Dans sa recherche, Vassiliadou (2008) adopte une démarche qui est proche de la nôtre. Elle travaille sur quatre marqueurs de reformulation *c'est-à-dire*, *à savoir*, *autrement dit*, *en d'autres termes*, dans l'objectif de relever les rapports entre ces marqueurs et également de mettre en évidence leurs différentes contraintes d'emploi. Ces quatre marqueurs sont regroupés, car ils relient, selon Vassiliadou, des « propositions équationnelles » sur la base d'équivalence sémantique. Les rapports établis entre les énoncés grâce aux marqueurs de reformulation **p** et **q** figurent sur deux plans, une équivalence métalinguistique et une équivalence qui relève de divers mécanismes comme des équivalences par métaphores, métonymies, synecdoques. Si le marqueur assure une équivalence métalinguistique qui relève de la synonymie lexicale et de la paraphrase linguistique, il porte soit sur le mot lui-même, soit sur le référent désigné par le mot, « que la référence soit actuelle, en discours, ou virtuelle, en langue » (*ibid.* : 47). Elle n'enregistre dans ce cas aucun changement de valeur de vérité de la proposition. Si le marqueur amène une équivalence qui relève des différents mécanismes, il introduit une équivalence d'identification, de localisation, de catégorisation, de qualification ou d'élaboration.

Outre ces régularités, Vassiliadou (*ibid.*) distingue l'emploi de ces marqueurs par des tests de compatibilité en prenant comme référence le marqueur *c'est-à-dire* – le marqueur de reformulation le plus prototypique selon elle. Ces tests révèlent des cas où les quatre marqueurs peuvent commuter, ainsi que des emplois qui spécifient un marqueur parmi d'autres. Ces différences sont dues d'une part, au sémantisme de chaque marqueur, et d'autre part à la structure syntaxique. Par exemple, *autrement dit* et *en d'autres termes*, par leur sémantisme, ne peuvent pas reformuler ce qui n'a pas été abordé précédemment. L'étude syntaxique et sémantico-pragmatique proposée nous permet de souligner de nouveau l'importance de l'analyse de la portée syntaxique des marqueurs de reformulation. Grâce à la portée, l'auteur peut relever les corrélations entre la séquence **p** et **q** qui sont d'ordre métalinguistique ou référentiel. Cette étude qui porte sur les quatre marqueurs de reformulation montre également l'intérêt d'une comparaison contrastée des MD pour mettre en relief les spécificités d'usage de chacun.

a. Définition

D'après le TLFi, *c'est-à-dire (que)* est utilisé pour introduire une information complémentaire ou rectificative. Si la définition du TLFi souligne surtout le rôle que joue ce marqueur dans l'énoncé, Khatchatourian (2008) tente d'expliquer la signification de *c'est-à-dire* par le sémantisme de son constituant dont le verbe *dire*, celui-ci désigne que « le locuteur extériorise (rend public) ce qu'il a à dire (sa représentation) en lui donnant une forme, *c'est-à-dire* en le mettant en mots » (*ibid.* : 32).

b. Portée syntaxique

En ce qui concerne la portée de ces marqueurs, *c'est-à-dire* peut avoir à la fois une portée globale et une portée locale. Il a une portée locale (**P_L**), quand *c'est-à-dire* porte sur un syntagme adjectival [66], un syntagme nominal [67], un syntagme verbal [68][69], ou un syntagme prépositionnel [70]. Quand *c'est-à-dire* porte sur une proposition, il a une portée globale (**P_G**) [71][72]. Dans les exemples suivants, la portée de *c'est-à-dire* est soulignée.

- [66] Tertullien ajoute que le vêtement « est entré » dans le Latium avec le mot, indiquant par là l'origine étrangère (**c'est-à-dire grecque**) du pallium, sans aucune précision chronologique. [his-art-127-body]
- [67] Il s'agit en fait d'appréhender dans quelle mesure le choix de l'opérateur du réseau (assurance maladie, mutuelle, assurance privée, professionnel de santé) favorisera une certaine adéquation entre l'efficacité, **c'est-à-dire la meilleure utilisation des ressources disponibles au sens de prestations de soins de qualité correspondant aux besoins des usagers**, et l'équité au sens d'égalité dans l'accès aux soins. [eco-art-254-introduction]
- [68] À l'inverse, la figure héroïque du flic dévoué qui prend à bras-le-corps tous les maux de la société dilue le rôle du policier dans celui du « travailleur social », et partant elle entérine l'idée selon laquelle la police fait « tout et n'importe quoi », **c'est-à-dire n'a plus de finalités assignées**. [infcom-art-5-body]
- [69] Mais il ne déclarait pas avoir exercé pendant dix ans, à compter de 1948, la fonction de chargé des relations publiques de la chambre syndicale du pétrole, **c'est-à-dire avoir été un « patronal »** (Richard (G.), « Industriels et partis de droite en France : essai d'approche globale », in Anceau (É.) et al., *Industrie et politique...*, op. cit. ; Dard (O.), Richard (G.), dir., *Les permanents patronaux*. [spo-art-433-notes]
- [70] La mobilité spatiale est ici considérée dans son sens large, **c'est-à-dire comme un mouvement qui s'exprime par un déplacement de l'individu dans l'espace, comme un mode de pratique de l'espace**. [geo-art-6-body]
- [71] Or, à l'exception de I2, qui conduit à ne l'aider que d'un montant suffisant pour le rendre indifférent entre la formation d'une fédération et l'autarcie, les institutions de négociation intergouvernementale I3 et I4 lui sont favorables dans toutes les circonstances (**c'est-à-dire quelle que soit la valeur de u**), à la fois pour les valeurs de u où elles rendent possible la formation d'une fédération qui serait impossible avec I1 et I2, et pour les valeurs de u telles qu'une fédération serait formée quelle que soit l'institution. [eco-art-293-body]

- [72] Ce paradoxe est lié au fait que l'égalitarisme numérique ne fait en fin de compte que reproduire l'inégalité de statut réelle entre membres et non- membres du parti : si chaque « coopain » peut participer aux actions du parti, seuls pourront y faire valoir leurs opinions ceux qui y seront statutairement intégrés, **c'est-à-dire** qui disposeront du titre officiel d'adhérents. [spo-art-340-body]

En revanche, *c'est-à-dire que*, en raison de sa nature syntaxique (plus précisément à la combinaison avec la conjonction *que*) ne peut porter que sur une proposition [73].

- [73] La vitesse était déterminée en utilisant une ligne de loch, telle que décrite dans l'ouvrage de Norie, mais le véritable cap du navire ne pouvait être estimé à partir du cap déterminé par le seul compas, car la force du vent faisait « dériver » le navire, **c'est-à-dire qu'**une partie du mouvement se faisait dans la même direction que le vent et à un degré déterminé par la force de celui-ci. [his-art-152-body]

Quelle que soit la portée de *c'est-à-dire*, nous constatons qu'il existe une certaine équivalence de la nature syntaxique entre le contexte gauche et la portée du marqueur ; cela veut dire que quand le contexte gauche est un syntagme nominal ou verbal ou adjectival, la portée de *c'est-à-dire* est également un syntagme nominal [67] ou verbal [68] ou adjectival [66]. Quand le contexte gauche est une proposition, la portée du marqueur est également une proposition [72]. Par exemple en [73], **p** = « la force du vent faisait “dériver” le navire », **q** = « une partie du mouvement se faisait dans la même direction que le vent et à un degré déterminé par la force de celui-ci ».

c. Position syntaxique

Nous distinguons *c'est-à-dire* en tant que locution adverbiale, et *c'est-à-dire que* comme locution conjonctive. Par rapport à la séquence **q**, comme nous l'avons remarqué pour les marqueurs de reformulation, *c'est-à-dire* et *c'est-à-dire que* sont toujours suivis d'une séquence reformulatrice, et ils se trouvent toujours en position initiale [74][75]. *C'est-à-dire* ne peut pas se trouver en position médiane ni en position finale. *C'est-à-dire* est en général précédé d'un signe typographique (le plus souvent une virgule). Au niveau phrastique, nous observons très peu de cas où *c'est-à-dire* ou *c'est-à-dire que* se trouvent en tête de phrase (6 occurrences de *c'est-à-dire*, 7 occurrences de *c'est-à-dire que*).

- [74] La répétition de ce rituel produit une certaine confiance parce qu'il témoigne d'une certaine stabilité : elle inscrit cette rencontre avec l'autre à peine identifié dans les limites d'un monde connu : elle contribue à ce que Giddens (1987) appelle la « sécurité ontologique », **c'est-à-dire** le sentiment que le monde est ordonné et stable, qu'il va de soi. [soc-art-59-body]
- [75] Comme le souligne Passeron, « discours statistique et le discours sociologique diffèrent intrinsèquement par leur pertinence empirique, **c'est-à-dire qu'**ils ne sont pas justiciables, pour trancher de la vérité ou de la fausseté de leurs énoncés, de

procédures de contrôle référant leurs expériences de dernière instance à la même réalité empirique » 1991, 113). [sed-art-8-notes]

d. Valeurs sémantiques

Nous distinguons quatre valeurs principales de *c'est-à-dire* et *c'est-à-dire que* dans l'énoncé : la valeur métalinguistique, la valeur interprétative, la valeur énumérative et la valeur conclusive.

D1. Valeurs métalinguistiques

C'est-à-dire peut avoir une valeur métalinguistique. Le rapport entre **p** et *c'est-à-dire* **q** est établi sur la relation synonymique. La séquence **q** apporte une explication, une précision, une définition, une dénomination. Quand *c'est-à-dire* sert à introduire la définition, **q** peut consister en une définition classique [76]. Il s'agit d'une signification indépendante de la situation de communication, ou bien d'une définition choisie par l'auteur, qui correspond donc au contexte d'énonciation [77].

[76] Ainsi la forme verbale (ils) « donnaient » est un mot composé d'un lexème (« donn- ») et de deux morphèmes écrits (« -ai- » et « -ent »). Stricto sensu, la morphographie d'une écriture, **c'est-à-dire**, l'écriture des formes grammaticales, est une composante de la sémiographie (pour un développement détaillé, se reporter à Jaffré, 2001), mais dans le cadre de cet article, la distinction entre les deux termes sémiographie et morphographie paraît secondaire dans la mesure où leurs relations avec la notion d'optimalité sont comparables. [lin-art-99-body]

Dans [76], l'auteur donne la définition exacte de *morphographie* dans la *morphographie d'une écriture* ; en [77], *c'est-à-dire* introduit un seul état de chose, ou une seule signification de l'énoncé **q**.

[77] Selon E.C. Hugues (Hugues E.C., 1996) une profession est caractérisée par le fait que l'exercice de ses activités nécessite l'obtention d'une licence, **c'est-à-dire** d'une permission légale d'exercer un certain travail interdit aux autres individus qui n'en sont pas détenteurs (ce qui peut correspondre, pour ce qui concerne les enseignants par exemple, à l'obtention du CRPE, CAPES, CAPET, CAPLP, agrégation, etc. [sed-art-124-body]

Dans l'exemple [77], **q** (dans *c'est-à-dire* **q**) est seulement une des définitions normées du mot *licence*. Le terme *licence* peut référer soit au diplôme national sanctionnant les trois premières années d'études du premier cycle de l'enseignement supérieur, soit au document administratif attaché à un véhicule routier, soit au permis d'exercer une activité, etc. Le marqueur *c'est-à-dire* introduit dans ce cas la définition du terme *licence* dans ce contexte et permet au locuteur de délimiter le référent de la *licence* et de faciliter le travail d'interprétation du lecteur.

C'est-à-dire peut également servir à introduire une définition pour un terme en langue étrangère.

- [78] Le social est devenu une réelle alternative au chômage, favorisant l'émergence d'une welfare class, **c'est-à-dire** un groupe social employé directement par les politiques publiques et qui en tire les principaux bénéfices. [spo-art-153-body]

L'énoncé **q** *un groupe social employé directement par les politiques publiques* est la définition du terme en anglais *welfare class*. Dans l'exemple [78], comme le terme en anglais n'est pas mis entre guillemets, nous pouvons en déduire qu'il fait partie de la terminologie de la discipline, en l'occurrence les sciences politiques. L'introduction de *c'est-à-dire* permet au locuteur d'explicitier l'énoncé pour le lecteur.

La définition introduite par *c'est-à-dire* (*c'est-à-dire que*) peut correspondre à la citation d'un autre auteur [79][80]. Dans ces exemples, les énoncés **p** dans ces deux exemples *effet domino* et *ensemble show* peuvent avoir différentes définitions. Pourtant, les citations correspondantes consistent à donner les significations que le locuteur veut aborder dans ce contexte précis.

- [79] Ce type de configuration spatiale augmente la probabilité d'un effet domino, **c'est-à-dire** « la survenance en cascades de dommages et de perturbations suite à un premier choc, au sein de l'enjeu touché, mais aussi dans son environnement (entendu au sens large d'environnement socio-économique) » (Provitolo, 2007 ; Gleyze, Reghezza, 2007). [geo-art-111-body]
- [80] Une autre innovation majeure de la série tient à sa distribution élargie, qui donne lieu à l'invention d'un nouveau genre, qualifié de « ensemble show », **c'est-à-dire** une « série mettant en scène un nombre important de personnages réguliers » (Winckler, 1985, p 23). [infcom-art-131-body]

D2. Les valeurs interprétatives

C'est-à-dire peut aussi introduire une valeur interprétative. Le locuteur peut proposer sa propre interprétation de **p**, d'où l'importance de préciser ce qu'il entend dans le contexte choisi [81][82].

- [81] Les résultats convergent pour suggérer que la conscience morphosyntaxique des apprenants, **c'est-à-dire**, leur capacité à prendre conscience des unités morphosyntaxiques et à les manipuler délibérément, se centre sur la forme des verbes français au détriment des conditions d'usage du prétérit en anglais. [lin-art-105-body]
- [82] Les deux composantes du traitement de la langue écrite, lire et orthographier sont en général considérées comme des activités « en miroir », **c'est-à-dire**, régies par les mêmes processus : d'une part les procédures alphabétiques, c'est-à-dire l'application des correspondances phonographémiques, et d'autre part les procédures lexicales avec la récupération automatique des représentations orthographiques dans le lexique mental (Sanguin Bruckert et Bruckert, 2001). [psy-art-109-body]

Dans l'exemple [81], le locuteur explicite dans **q** ce qu'il voulait exprimer par la *conscience des unités morphosyntaxiques*. Dans l'exemple [82], *c'est-à-dire* introduit les interprétations de l'auteur sur le terme **p** *en miroir*. L'auteur a choisi de mettre *en miroir* entre guillemets, car il lui prête une signification spécifique. *C'est-à-dire* détermine ici la seule acception du terme dans ce contexte. Nous soulignons dans cet exemple [82] la concordance grammaticale de l'énoncé **q** quand celui-ci vise à expliciter le sens de l'énoncé **p** *en miroir* et joue le rôle de complément des *activités*.

Nous constatons un autre emploi de *c'est-à-dire* quand celui-ci introduit un hyperonyme [83]. *C'est-à-dire* instaure dans cet exemple une relation d'hyponymie entre **p** et **q**. La séquence **q** correspond à quelque chose de plus général tandis que **p** renvoie aux éléments plus concrets *professions médicales, puis les métiers du sport et des arts*.

- [83] Les métiers qui portent la marque d'un onirisme social sont très orientés : ils concernent pour l'essentiel les professions médicales, puis les métiers du sport et des arts, **c'est-à-dire** des métiers nobles et prestigieux, clairement inaccessibles. [soc-art-307-body]

La reformulation établie par *c'est-à-dire* peut être employée pour interpréter l'énoncé **p**. Elle correspond à une conception subjective du locuteur de l'état de choses. *C'est-à-dire* apporte les interprétations du locuteur, et également d'autres informations qui permettent de situer le contexte [84][85][86].

- [84] À titre d'exemple, la rubrique d'un petit forum d'initiés concernant la série Breaking Bad, forte de nombreux échanges entre une douzaine de fans durant la deuxième saison présentée aux États-Unis, cesse brusquement toute activité pendant l'intersaison, de juin 2009 à mars 2010, pour reprendre précisément le 22 mars, **c'est-à-dire** le jour suivant la diffusion américaine du premier épisode de la troisième saison. [infcom-art-21-body]
- [85] Dès l'aube du XIV^e siècle, **c'est-à-dire** au moment où les testateurs resserrent les cordons de leur bourse, nombre de fidèles qui continuent de favoriser les institutions charitables par le biais de leurs administrateurs ne désignent plus pour ce faire que les recteurs de l'hôpital Saint-Esprit presque tous d'origine mercantile, comme s'ils cherchaient à garantir, sinon à contrôler (et non aliter) l'exécution réelle de leurs vœux [his-art-63-body]
- [86] Ce n'est sans doute pas un hasard si le Front national obtient au début des années 1990 **c'est-à-dire** à un moment où la cité est devenue aux yeux de ses premiers habitants un quartier de relégation que la présence « étrangère » conduit à stigmatiser ses meilleurs résultats électoraux : jusqu'à 45 % des suffrages au second tour des élections législatives de 1993. [spo-art-160-body]

Dans ces exemples, les auteurs ont besoin de préciser l'évènement correspondant à chaque notion temporelle. L'utilisation de ce marqueur permet donc à l'auteur de mettre le projecteur sur l'évènement afin d'éclairer le lecteur sur ce qui s'est passé dans un contexte précis. En particulier, nous relevons une autre valeur d'emploi de *c'est-à-dire* qu'on mettra

parmi les valeurs interprétatives : il s'agit de la valeur de renvoi. *C'est-à-dire* renvoie dans ce cas directement à un élément évoqué précédemment dans le texte [87]. À la différence des valeurs de définition, *c'est-à-dire* y introduit également une explication de l'auteur.

- [87] En ce qui concerne la solution (3), **c'est-à-dire** l'émission de certificats de dépôts, il s'agit de savoir si la demande de certificats de dépôts est parfaitement élastique au taux du marché, **c'est-à-dire** si les banques peuvent émettre autant de certificats de dépôts qu'elles le veulent sans payer de prime. [eco-art-572-body]

Dans l'exemple [87], la solution en question *la solution (3)* fait déjà l'objet d'analyses antérieures. Le premier *c'est-à-dire* consiste en un rappel et introduit une suite logique d'autres analyses. Dans cet exemple, *c'est-à-dire* est utilisé deux fois dans la même phrase. Si le premier *c'est-à-dire* sert à référer à ce qui a été dit, le deuxième *c'est-à-dire* a une valeur interprétative quand l'auteur cherche à expliciter l'énoncé **p** *si les banques peuvent émettre autant de certificats de dépôts qu'elles le veulent sans payer de prime*.

D3. Valeur énumérative

C'est-à-dire peut avoir une valeur énumérative [88][89][90]. L'utilisation de *c'est-à-dire* est possible dans ces exemples, car la séquence **q** introduit des connaissances qui se trouvent dans le répertoire du lecteur ou sont acquises par son expérience, ce qui peut être interprété par l'usage du déterminant défini, comme *le* dans [90]. Ce fonctionnement de *c'est-à-dire* est un des critères qui sert à distinguer ce marqueur de *à savoir*, comme on le verra par la suite.

- [88] Le nombre d'internautes s'élevait en janvier 2008 à 210 millions de personnes (hors régions administratives spéciales, **c'est-à-dire** Macao, Hong Kong et Taïwan) ; la progression est de 53 % sur l'année 2007, soit de 100 internautes à la minute ; la Chine est ainsi devenue fin 2007 le premier pays au monde pour le nombre d'internautes. [geo-art-34-body]
- [89] Son effectif est constitué des experts déjà investis dans les organismes nationaux, **c'est-à-dire** des agronomes et des juristes qui se réclament de travaux publiés par les géographes. [spo-art-51-body]
- [90] 'Il [l'officier] doit, au moment de sa montée à bord, tenir un Journal... et faire attention d'y noter tous les événements, **c'est-à-dire** l'endroit où le navire se trouve à midi : les changements de vent et de temps... les observations sur les lieux inconnus ; et en général, toute circonstance qui concerne le navire... il doit envoyer une copie de son journal, à l'heure dite, au secrétaire de l'Amirauté'. [his-art-152-body]

Dans l'exemple [88], l'énoncé **p** *régions administratives spéciales* correspond à une façon plus subjective d'exprimer l'état de choses. Dans l'exemple [89], l'énoncé **q** délimite ce que l'auteur veut aborder dans l'énoncé **p** et dans l'exemple [90], le MD permet à l'auteur de détailler une succession d'événements que l'officier doit noter dans le journal. Dans les

exemples [88][89], *c'est-à-dire* instaure en quelque sorte une relation d'hyponymie en allant de quelque chose plus général (dans **p**) à quelque chose de plus spécifique ou plus précis (dans **q**).

D4. Valeur résumante

Enfin, l'emploi de *c'est-à-dire* assure une valeur résumante, même si cet usage n'est pas fréquent dans notre corpus [91].

- [91] La photographie n'est nullement un langage, si ce n'est précisément par l'abus de langage qui la transforme en discours avec des mots, une grammaire et une syntaxe, **c'est-à-dire** en ce qui est aphotographique. [ant-art-265-body]

La séquence **q** résume en fait ce qui a été dit dans la séquence **p** par d'autres mots. *C'est-à-dire* a une valeur résumante et exprime dans une certaine mesure l'avis catégorique de l'auteur de *ce qui est aphotographique*. *C'est-à-dire* ne sert pas à introduire des arguments de l'auteur comme d'autres marqueurs (*autrement dit*, *en d'autres termes*), comme nous le verrons par la suite.

Nous allons maintenant comparer la taille de deux séquences **p** et **q** pour ensuite observer le lien entre cet aspect syntaxique et les valeurs sémantiques du MD. Nous distinguons trois cas de figure : 1/ la taille de **p** et **q** est équivalente ; 2/ la taille de **p** est supérieure à celle de **q** ; et 3/ la taille de **p** est inférieure à celle de **q**. (Cf. 5.2.4, page 117). Quand **p** et **q** sont de même taille, *c'est-à-dire* que ces deux séquences sont de même structure syntaxique, *c'est-à-dire* peut assurer une valeur métalinguistique [92], ou une valeur explicative [93], ou bien une valeur interprétative [94].

- [92] Pour Sraffa et les néocambridgiens, le conflit pour la répartition, *c'est-à-dire* le partage du surplus entre salaires et profits, conditionne la valorisation même de ce surplus, *c'est-à-dire* la détermination des prix de production (et de reproduction) de l'économie nommément reconnue et analysée comme 'capitaliste', étant donné par ailleurs la matrice des coefficients techniques de production ; dans cette analyse, seuls les 'rentiers', **c'est-à-dire** les propriétaires des ressources naturelles (la 'terre', *lato sensu*), peuvent arbitrer ce conflit en demandant des prélèvements, en nature ou en valeur, sur le surproduit. [eco-art-275-introduction]
- [93] Par ailleurs, quel que soit le lieu de production, en ambulatoire ou en établissement, cette production bénéficiera de la stabilité qui accompagne l'intégration (par opposition à la flexibilité qui accompagne la désintégration) : dans une organisation coordonnant les soins pour un plus grand nombre de patients, les professionnels de santé acquièrent davantage d'expérience et donc une plus grande compétence qui peut s'analyser ici comme un "effet club", **c'est-à-dire** une externalité liée au fait que la

valeur de l'appartenance à l'organisation augmente lorsque le nombre d'utilisateurs s'accroît. [eco-art-254-body]

- [94] La plupart des études chez l'adolescent ont été transversales, **c'est-à-dire** peu propices à distinguer, dans la relation entre stress et coping, ce qui était cause ou conséquence. [psy-art-178-introduction]

Nous soulignons deux cas où la taille de ces énoncés **p** et **q** sont différents, ce qui entraîne une différence de valeur d'emploi. Quand la taille de **p** est supérieure à celle de **q**, *c'est-à-dire* introduit une valeur résumante [91], ou une valeur métalinguistique par dénomination [95]. Quand la taille de **p** est inférieure à celle de **q**, *c'est-à-dire* porte une valeur énumérative [96].

- [95] Il ne s'agit pas de s'intéresser aux savoirs scientifiques, aux comptes rendus exhaustifs des pratiques, à la présentation exacte de projets divers, etc., mais à tout ce qui constitue les contenus des diverses communications non réglées par l'organisation des enseignements et en particulier les opinions, les connaissances, les attitudes, les jugements et les croyances des personnes interrogées vis-à-vis de l'informatique, **c'est-à-dire** aux représentations sociales de cet objet. [sed-com -208-body]
- [96] Les plus anciens de ses informateurs pensaient que la manière de chanter des jeunes n'était plus celle dont ils étaient eux-mêmes, et que l'accélération du rythme qu'ils percevaient venait de la fréquentation des Piémontais (**c'est-à-dire** des bûcherons et des ouvriers résidant à Tende avant la guerre) qui chantaient plus rapidement que les Tendasques. [ant-art-222-body]

Nous pouvons résumer les propriétés syntaxiques et sémantiques de *c'est-à-dire* dans le Tableau 5.1 : Récapitulatif des propriétés syntaxiques et sémantiques de *c'est-à-dire*.

Paramètres	Propriétés
Position	Initiale. Au niveau phrastique, il se trouve en position postverbale. Quand il porte sur une proposition, il est placé en tête de phrase.
Portée	<p>Locale. Le contexte gauche p a la même structure syntaxique que la portée q.</p> <p>Globale, uniquement pour <i>c'est-à-dire que</i>.</p> <div style="text-align: center;"> <p>The diagram illustrates the syntactic structure of 'c'est-à-dire' in two contexts. In the local scope, 'c'est-à-dire' is positioned between two phrases. Each phrase is a bracketed group containing four syntagmas: nominal, verbal, prepositional, and adjectival. The left phrase is labeled 'p' and the right phrase is labeled 'q' and 'phrase'. In the global scope, 'c'est-à-dire que' is positioned between two propositions, each labeled 'p' and 'q' respectively.</p> </div>
Valeurs sémantiques	<p>✓ <i>Valeurs métalinguistiques</i> : <i>c'est-à-dire</i> peut servir à introduire une définition, une dénomination, une explication. Il établit un rapport synonymique entre la séquence p et q.</p> <p>[]_p c'est-à-dire_{V.METALINGUISTIQUE} []_q</p> <p><p> = <q></p> <p>✓ <i>Valeurs interprétatives</i> : <i>c'est-à-dire</i> apporte de propres interprétations du locuteur (par des conceptions des choses, par renvoi dans le texte, etc.).</p> <p>[]_p c'est-à-dire_{V.INTERPRETATIVE} []_q</p> <p><p> = <q></p> <p>✓ <i>Valeurs énumératives</i> : <i>c'est-à-dire</i> introduit une énumération. L'utilisation du déterminant défini indique que l'information est connue par le lecteur par ses propres connaissances ou par les informations dans le texte.</p> <p>[]_p c'est-à-dire_{V.ENUMERATIVE} []_q</p> <p><p><q></p>

	<p>✓ Valeurs résumante : <i>c'est-à-dire</i> introduit une idée synthétique.</p> <p>[]_p c'est-à-dire[V.RESUMANTE] []_q</p> <p><p>><q></p>
--	--

Tableau 5.1 : Récapitulatif des propriétés syntaxiques et sémantiques de *c'est-à-dire*

Grâce aux exemples analysés, nous pouvons tirer quelques conclusions sur les valeurs d'emploi de *c'est-à-dire*. Premièrement, on constate une équivalence syntaxique entre **p** et **q** : en général, la séquence **q** reprenant la même structure syntaxique que **p**. Deuxièmement, nous pouvons relever des corrélations fortes entre les valeurs sémantiques de *c'est-à-dire* et la taille de sa portée **q** et son contexte gauche **p**.

Pour résumer, *c'est-à-dire* peut prendre sous sa portée quasiment tous les constituants (mot, syntagme nominal, adjectival, verbal et prépositionnel, proposition). Ce marqueur peut assumer quatre valeurs sémantiques principales : la valeur métalinguistique, la valeur interprétative, la valeur énumérative et la valeur conclusive. Dans le Tableau 5.1, figurent d'une manière concise différentes valeurs sémantiques de *c'est-à-dire*, ainsi que la taille de la séquence **p** en comparaison avec celle de la séquence **q** dans chaque cas correspondant. Ce tableau n'englobe peut-être pas tous les cas de figure sur le rapport syntaxique entre **p** et **q**, mais il permet dans une certaine mesure de mettre en évidence des régularités récurrentes entre ces séquences, ce qui est indispensable dans l'enseignement/apprentissage de ces éléments linguistiques.

5.3.2. À savoir

a. Définition

À *savoir*⁴⁰ connaît une fréquence beaucoup plus faible que *c'est-à-dire* (139 occurrences). À *savoir* est souvent considéré comme un synonyme de *c'est-à-dire*. Certes, dans quelques cas, *c'est-à-dire* peut être utilisé à la place de à *savoir* pour exprimer une équivalence. Dans d'autres cas, l'utilisation de *c'est-à-dire* au lieu de à *savoir* devient moins

⁴⁰ À *savoir* est défini comme une locution conjonctive ou adverbiale, il introduit une explication ou un développement (d'après le TLFi).

naturel et même impossible. Nous allons analyser l'utilisation de *à savoir* dans notre corpus afin de faire ressortir ses propriétés sémantiques et discursives. Nous retenons dans cette recherche uniquement les occurrences de *à savoir* qui jouent le rôle d'une locution adverbiale⁴¹.

b. Portée syntaxique

À savoir porte essentiellement sur un syntagme nominal [97][98], même s'il peut moins fréquemment également porter sur un syntagme prépositionnel [99] ou sur un syntagme verbal [100]. Ce marqueur a, en général, une portée locale (**P_L**) et nous relevons très peu d'occurrences de *à savoir* portant sur une proposition [104].

- [97] Selon les néostructuralistes, l'accent mis sur les équilibres macroéconomiques (et notamment l'atteinte de l'équilibre commercial par la diminution de l'absorption interne) a conduit à négliger ce qui constitue le facteur d'impulsion spécifique de la formation du capital, **à savoir** le niveau de l'activité économique : les réductions sensibles de l'investissement net et de l'usage de la capacité de production installée ne faisaient qu'accompagner la baisse du niveau d'activité constatée au cours de cette décennie. [eco-art-275-body]
- [98] Examinons maintenant les théories issues du paradigme B, **à savoir** celles qui ont toutes comme point commun une définition compositionnelle de la signification. [lin-art-509-body]
- [99] Cette dernière se constitue autour d'un nodal point, **à savoir** autour du point central d'un signifiant fixe, d'un terme ou d'un thème par rapport auquel l'ensemble des significations s'organise de manière contingente en réseaux d'opposition et d'équivalence mouvants (Laclau et Mouffe, 2009 [1985]). [geo-art-282-body]
- [100] Ils proposent un point de vue communicationnel sur ces objets, **à savoir** décrypter les rôles et légitimités d'auteurs à travers une analyse énonciative (qui parle, de quoi, comment l'exprime-t-il), médiatique (par quel média, sur quel support) et éditoriale (à quel moment, en direction de qui, avec quel type de traduction). [infcom-art-311-body]

En observant le contexte gauche **p** de *à savoir*, nous constatons que la séquence **p** est rarement de nature propositionnelle, mais elle correspond plutôt à un syntagme nominal. Il existe généralement une certaine équivalence syntaxique entre **p** et **q** : *à savoir* relie deux syntagmes nominaux, verbaux ou prépositionnels.

⁴¹ Toutes les occurrences de *à savoir* qui apparaissent dans des structures comme *il reste/Reste à savoir*, ou *V + à savoir* (comme *chercher/revenir/demander/consister à savoir*) sont exclues. Dans ces cas, *à savoir* fait partie de la structure de la phrase et ne consiste pas en une séquence lexicalisée à fonction discursive.

c. Position syntaxique

À *savoir* se trouve uniquement en position initiale [101][102]. Ce marqueur est souvent précédé d'une virgule [103], et peut être suivi par deux points [104]. Nous n'avons trouvé aucune occurrence de *à savoir* utilisé comme locution adverbiale en tête de phrase.

- [101] Il permettait également d'associer les deux noyaux historiques de la ville, **à savoir** l'implantation portuaire orientale (parallèle à la Breedstraat) et la partie longtemps demeurée agricole à l'ouest. [his-art-363-body]
- [102] Par contre, l'affirmation verbale, **à savoir** l'utilisation du non- sémantique, implique un nouveau niveau d'autonomie, dans lequel l'enfant devient conscient de l'autre « et de » soi », et s'observe dès la fin de la deuxième année (Crockenberg et Litman, 1990). [psy-art-382-body]
- [103] Les 3 pays étudiés, **à savoir** la France, l'Espagne et l'Italie, se sont engagés dans l'élaboration d'une politique paysagère pour partie commune, notamment dans le cadre de la CEP, tout en présentant des approches spécifiques. [geo-art-393-body]
- [104] La deuxième hypothèse dit l'inverse, **à savoir** : plus le thème est inscrit dans des dispositifs émanant d'institutions spécifiquement diffuseuses de nouvelles, plus il circulera dans des dispositifs médiatiques à l'initiative des institutions qui ont des objectifs divers. [infcom-art-366-body]

d. Valeurs sémantiques

D1. Valeur énumérative

Une des valeurs les plus fréquentes de *à savoir* dans notre corpus est la valeur énumérative [105][106][107].

- [105] Outre la vigilance qu'elles suggèrent à l'égard des constats avancés dans la littérature sur la judiciarisation, ces considérations critiques permettent de comprendre les bilans contrastés que cette littérature révèle pour les deux axes thématiques que nous avons choisi d'évoquer dans cet article, **à savoir** le rôle de la justice dans la fabrique des politiques publiques d'une part et dans le fonctionnement démocratique d'autre part. [soc-art-35-body]
- [106] Entre 1998 et 2008, 176 articles sont répertoriés sous le thème « mécanisme » dans les quatre revues majeures de sociologie (**à savoir** l'American Journal of Sociology, l'Annual Review of Sociology, l'American Sociological Review et Social Forces) contre 56 entre 1956 et 1997 (source : ssci). [soc-art-253-notes]
- [107] On peut toutefois rejoindre Alain Caillé (2007 : 196) lorsqu'il s'interroge sur les rapports ou combinaisons que peuvent prendre entre eux les trois principes de reconnaissance distingués par Axel Honneth (2007), **à savoir** la sphère de l'intimité (confiance en soi, amour), la sphère juridique (respect de soi) et la sphère de la solidarité (estime de soi). [infcom-art-344-body]

Dans ces exemples, *à savoir* sert à introduire une série d'éléments. L'énumération précédée de *à savoir* permet d'explicitier ce qui a été abordé dans **p**. C'est pourquoi *à savoir* apporte toujours de nouvelles informations à l'énoncé, ce qui peut s'expliquer par le

sémantisme des constituants du marqueur « à + savoir » que l'on pourrait gloser en « ce qu'il faut savoir/connaître ». Par exemple, dans l'exemple [105], la phrase sera incomplète si l'on s'arrête à *dans cet article*, car le lecteur s'attend à apprendre quels sont les *deux axes thématiques* évoqués dans la séquence **p**. À *savoir* apporte en quelque sorte l'information manquante dans la séquence **p**. Dans l'exemple [106], sans la séquence **q** introduite par *à savoir*, le locuteur posera la question « Quelles sont les quatre revues majeures de sociologie » ? Ou bien « Quels sont les trois principes de reconnaissance distingués par Axel Honneth ? » comme dans l'exemple [107]. Nous remarquons souvent l'utilisation de numéros cardinaux dans le contexte gauche avant l'introduction d'une énumération.

Dans les exemples [105][106][107], *à savoir* peut être remplacé par *c'est-à-dire* uniquement si l'on peut supposer que le lecteur a acquis déjà ces connaissances par ses propres expériences ou par des précisions précédentes dans le texte. L'emploi de *c'est-à-dire* dans les exemples [108][109] devient moins naturel, car l'énumération correspond à la séquence **q**, apporte de nouvelles informations, de nouveaux savoirs au lecteur.

- [108] Avant de considérer les différents tests de la théorie des anticipations, il est nécessaire de valider au préalable les hypothèses sous-jacentes à la représentation RVAR, à savoir d'une part, de la non- stationnarité des taux d'intérêt et, d'autre part, de la stationnarité des pentes des taux. [eco-art-95-body]
- [109] Cet article présente les trois grandes études menées lors de notre programme de recherche à savoir, deux sur le développement de stimuli adaptés à une population francophone (études 1 et 2) et une troisième menée au laboratoire pour tester en français la généralisabilité du second postulat de la théorie de la catégorisation à fondement émotionnel (i. e., réplique de Niedenthal et al., 1999, expérience 1). [psy-art-246-body]

Comme nous l'avons analysé ci-dessus, *c'est-à-dire* peut également introduire une valeur énumérative, mais *c'est-à-dire* et *à savoir* ne sont pas interchangeables dans tous les contextes. Certes, nous pouvons considérer à première vue que l'emploi de *à savoir* dans les exemples [108][109] est le même que dans les exemples [105], mais *à savoir* est remplaçable par *c'est-à-dire* uniquement dans [105]. En fait, comme nous l'avons vu précédemment, *c'est-à-dire* ne peut reformuler que ce qui a été dit dans **p**. Effectivement, par le temps du verbe dans [105], nous pouvons deviner que *les deux axes thématiques* ont été précédemment évoqués dans le texte, tandis que la séquence **q** introduite par *à savoir* dans [108][109] sert à catégoriser ce qui est dit dans la séquence **p** et cela s'effectue par une série d'éléments. Dans l'exemple [109], le locuteur précise les *trois grandes études menées* lors de leur programme

de recherche, comme dans l'exemple [108], les informations dans la séquence **q** n'ont pas été abordées dans le texte.

Si l'on considère l'énumération suivie par *c'est-à-dire* plutôt comme un rappel de ce qui a été abordé précédemment, à *savoir* introduit une énumération qui sert à spécifier la séquence **p**. Nous reprenons la définition de la spécification de Murat & Cartier-Bresson (1987), « la spécification permet le passage direct d'un ensemble à ses éléments constituants, sans reprise énonciative ni coréférence » (Murat & Cartier – Bresson, 1987, cité par Vassiliadou, 2008 : 47).

- [110] Cela s'explique bien entendu par ce que Foucault avait identifié comme étant l'une des caractéristiques des sociétés modernes, **à savoir** « l'importance croissante prise par le jeu de la norme aux dépens du système juridique de la loi. (Ce qui ne veut pas dire) que la loi s'efface ou que les institutions de justice tendent à disparaître ; mais que la loi fonctionne toujours davantage comme une norme, et que l'institution judiciaire s'intègre de plus en plus à un continuum d'appareils (médicaux, administratifs...) dont les fonctions sont surtout régulatrices » (Foucault, 1976, p 189 - 190). [soc-art-162-body]
- [111] Ce résultat apparaît d'autant plus élevé que notre échantillon est composé à 60 % de « nouveaux adhérents », **à savoir** d'adhérents entrés à l'UMP depuis 2005. [spo-art-73-notes]
- [112] Si le sujet n'a pas identifié le changement à la fin de la présentation, ce dernier se voit attribuer comme temps la durée totale de la présentation, **à savoir** 100 s pour la version rapide et 400 s pour la version lente. [psy-art-185-body]

Dans les exemples ci-dessus, la séquence **q** introduite par *à savoir* sert à spécifier le contenu de la séquence **p**. L'emploi de *c'est-à-dire* nous paraît impossible dans ces contextes. Dans l'exemple [111], l'auteur ne cherche pas à donner une nouvelle définition de la séquence **p**, mais à catégoriser, à préciser qui sont ces *nouveaux adhérents*. L'emploi de *à savoir* dans [110] et [112] a la même valeur. Il sert à spécifier la séquence **p**.

D2. Valeurs métalinguistiques

À *savoir* peut introduire une séquence à valeur métalinguistique. Il peut exprimer un rapport synonymique par des procédés de définition [113], par interprétation [116][117], par renvoi [114][115].

- [113] Les seules tentatives d'évaluation sont celles qui reposent sur des données génériques des actions de politique documentaire, soit évaluation proposée en amont, en tout début de projet, soit en aval, **à savoir** en toute fin pour tenter d'apprécier la part d'impact des actions - notamment de gestion - menées dans l'espace documentaire. [infcom-art-62-body]

- [114] Confronter ces deux dernières possibilités repose alors sur la seconde statistique, à savoir la mesure du changement ou de la variation due à la prise en compte de deux points de vue différents. [psy-art-194-body]
- [115] Si l'on veut réintroduire un raisonnement expérimental dans une telle étude, à savoir un raisonnement guidé par un schème causal (Berthelot, 1990), il faut contrôler les facteurs influents de manière à tester l'effet de la variable qui nous intéresse (en l'occurrence, l'internalité) toutes choses égales par ailleurs. [sed-art-387-body]
- [116] Imaginer ces effets opposés contribue à développer une posture réflexive à l'égard de son dispositif de recherche et de son activité d'écriture, tout en sachant que le lecteur lui-même donnera sa traduction et son interprétation des points de vue élaborés, ici, à partir de ce qui interpelle les atsem, à savoir un positionnement complexe dans un métier pouvant apparaître comme simple. [ant-art-183-body]
- [117] On peut repérer d'autres marqueurs de ségrégation dans l'interprétation faite par certains membres de l'équipe de l'une de mes erreurs que l'on peut, suivant Goffman, désigner de « fausse note », à savoir une rupture dans la définition commune et acceptée par tous de la situation [1973].

L'utilisation de *c'est-à-dire* est acceptable dans ces exemples, car on peut supposer que la séquence **q** introduit des connaissances précédemment évoquées dans le texte [114][115] ou faisant déjà partie du savoir du lecteur, comme la définition de la notion *en aval* dans [113] ou celle des *atsem* et de la *fausse note* dans [116][117]. Dans l'exemple [114], l'utilisation du déterminant défini dans la séquence **p** prouve que le locuteur peut avoir abordé cette *seconde statistique* dans son texte. À *savoir* peut ici être remplacé par *c'est-à-dire*, car le locuteur cherche à expliquer ce qu'il entend par la notion de *raisonnement expérimental* dans son étude.

- [118] Le constat général d'échec des Programmes d'Ajustement Structurel mis en œuvre durant les années 80 et 90 par les Institutions Financières Internationales signale non seulement que les mesures d'inspiration néolibérale sont incapables de répondre de manière adéquate aux difficultés spécifiques que rencontrent actuellement les pays en voie développement, mais indique aussi et surtout qu'il est nécessaire de recourir à une approche plus réaliste et plus convaincante des problèmes et des situations de sous-développement, à savoir l'approche néostructuraliste. [eco-art-275-summary]

Dans l'exemple [118], l'emploi de *à savoir* peut ressembler à une des valeurs sémantiques de *c'est-à-dire* que nous avons analysées, la valeur métalinguistique, quand celui-ci introduit une dénomination. Pourtant, l'emploi de *c'est-à-dire* ne peut pas fonctionner dans l'exemple [119][120]. Effectivement, dans ces exemples, la séquence reformulatrice **q** introduite par *à savoir* apporte une information nouvelle pour le lecteur, ce qui se manifeste par l'utilisation du déterminant indéfini et de l'adjectif *autre*.

- [119] Désormais les testateurs se soucient des prisonniers d'un autre ordre, à savoir ceux dont la gêne économique a réduit à l'emprisonnement pour insolvabilité : dans leur testament respectif daté de juin 1326, Cécile et son époux, le négociant Guillaume

Égide, désirent assurer l'entretien de prisonniers pour dettes, voire même financer leur liberté [his-art-63-body]

- [120] L'observation participante sur les tournages révèle une autre caractéristique de la scripte, **à savoir** une certaine capacité d'effacement. Incarnant la mémoire du tournage et représentant un repère central pour la fabrication du film, la scripte est souvent consultée par les réalisateur (trice) s, mais celles que j'ai observées se manifestaient en revanche très peu auprès de ces derniers. [ant-art-192-body]

Dans l'exemple [119], le locuteur développe ses arguments en ajoutant dans la séquence **q** une autre raison pour laquelle *les lecteurs se soucient des prisonniers*. L'emploi de *à savoir* dans l'exemple [120] est le même que celui dans l'exemple [119] quand la séquence **q** caractérise le contenu de la séquence **p**. L'auteur cherche en fait à mieux décrire les caractéristiques de *la scripte*.

Résumons ce qui est à retenir pour la description de l'utilisation de *à savoir*. Si *c'est-à-dire* relie en général deux séquences de taille équivalente, la taille de **p** et **q** reliées par *à savoir* est rarement égale. Quand elles le sont, il peut y avoir une valeur métalinguistique, et *à savoir* peut donc être remplacé par *c'est-à-dire* [114][115][116][117]. Quand la taille de **q** est supérieure à celle de **p**, l'auteur cherche à développer ses arguments et à apporter de nouvelles informations ; alors *à savoir* assume une valeur énumérative [121] ou une valeur de spécification [122]. La valeur énumérative paraît être plus fréquente pour l'utilisation de *à savoir*, ce qui permet de le distinguer de *c'est-à-dire*.

- [121] Trois dimensions sont particulièrement étudiées **à savoir** : la définition et délimitation du travail autour du traitement documentaire, la question de la production de ressources documentaires, les modes d'accompagnement mis en œuvre par les documentalistes afin d'explicitier le traitement et l'accès à l'information (sous forme de formations, de sensibilisation, de conseils, etc.). [infcom-art-62-body]
- [122] La crainte, au demeurant tout à fait légitime, de tomber dans le piège d'une herméneutique qui s'affranchirait des données de l'expérience et se saisirait d'elle - même objet de discours, ne devrait pas conduire à abandonner ce qui constitue une fonction essentielle du langage, **à savoir** la construction de significations, fussent elles révisables en fonction de l'évolution des mondes à décrire. [ant-art-41-body]

Si la taille de la portée de *à savoir* est réduite – la séquence **q** est inférieure à celle de **p** - alors *à savoir* a une valeur métalinguistique et introduit une dénomination [123][124].

- [123] L'emploi des termes et leur répétition (notamment le terme spectaculaire) dans les énoncés des ornithologues et gestionnaires de zones humides, dont la plus grande part sont la propriété du Conservatoire montre bien l'objectif recherché : donner à voir en provoquant certains phénomènes naturels par l'ingénierie écologique pour les rapprocher des spectateurs, **à savoir** le public. [geo-art-298-body]
- [124] Cette propriété commune, si elle existe, ne peut résulter que de ce qu'ils ont en commun et qui les distingue des plantes et de l'inorganique, **à savoir** le souffle divin. [ant-art-368-body]

Dans le Tableau 5.2 : Récapitulatif des propriétés syntaxiques et sémantiques de *à savoir*, figurent les propriétés syntaxiques et sémantiques principales de *à savoir*.

Paramètres	Propriétés
Position	Initiale. Au niveau phrastique, il se trouve en position postverbale.
Portée	<p>Locale. <i>À savoir</i> porte toujours sur les syntagmes.</p> $\left(\begin{array}{c} \text{Syntagme nominal} \\ \text{Syntagme verbal} \\ \text{Syntagme prépositionnel} \end{array} \right)_{\mathbf{p}} \text{ à savoir } \left(\begin{array}{c} \text{Syntagme nominal} \\ \text{Syntagme verbal} \\ \text{Syntagme prépositionnel} \end{array} \right)_{\mathbf{q}}$
Valeurs sémantiques	<p>✓ <i>Valeurs énumératives</i> : <i>à savoir</i> introduit une énumération qui tente de spécifier la séquence p. La taille de la séquence q est supérieure à celle de p.</p> <p>[]_p <i>à savoir</i> [V. ENUMERATIVE] []_q</p> <p><p><<q></p>
	<p>✓ <i>Valeurs interprétatives</i> : <i>à savoir</i> peut établir un rapport synonymique entre p et q pour introduire une définition, une dénomination, une explication. La taille de la séquence p est égale à celle de q.</p> <p>[]_p <i>à savoir</i> [V. METALINGUISTIQUE] []_q</p> <p><p> = <q></p>

Tableau 5.2 : Récapitulatif des propriétés syntaxiques et sémantiques de *à savoir*

En conclusion, *à savoir* peut disposer de deux valeurs sémantiques principales : énumérative et métalinguistique (Tableau 5.2). La valeur énumérative constitue un emploi spécifique de ce marqueur, mais *à savoir* peut dans quelques cas être remplacé par *c'est-à-dire* quand il dispose d'une valeur métalinguistique. En d'autres termes, le rapport entre **p** et **q** est établi sur un rapport d'équivalence par des procédés de définition ou d'explication.

5.3.3. Autrement dit

a. Définition

Autrement dit est défini comme une locution reprenant « sous une nouvelle formulation quelque chose qui vient d’être dit » (d’après le TLFi). La signification de *autrement dit* est liée au sémantisme de ses constituants : la séquence **q** introduite par *autrement dit* consiste à utiliser d’autres mots pour dire la même chose dans **p**. Pourtant, cette définition du TLFi ne précise pas différentes utilisations de ce marqueur, par exemple : à l’instar des marqueurs précédents, *autrement dit* peut également introduire une séquence à valeur explicative [125], mais il se caractérise aussi par d’autres usages que nous allons analyser par la suite.

- [125] Les analyses du MID participent d’un tournant accompagné par le pouvoir brejnévien dès 1966, lorsqu’une réunion du Politburo avait été spécialement convoquée pour discuter des « questions du travail idéologique » **autrement dit**, des insuffisances dans les méthodes du travail de propagande. [his-art-513-body]

b. Portée syntaxique

En raison de sa structure syntaxique⁴², *autrement dit* prend en général sous sa portée une phrase entière [126][127] et a portée globale (**P_G**). Quand *autrement dit* porte sur une proposition, son contexte gauche **p** est aussi de nature propositionnelle et peut éventuellement dépasser la phrase.

- [126] Ce que l’on regarde sur la photo ou sur l’écran n’est pas le « monde » mais un fragment saisi à un moment particulier d’une temporalité en devenir. Ce qui peut être connu, ce sont des aspects du réel (que l’on pourrait aussi bien nommer du graduel, du processuel ou de l’événementiel) ne pouvant être appréhendés qu’à partir d’une certaine perspective. Ce qui est impossible en photographie ou au cinéma, tous deux condamnés à la singularité concrète, ce qui leur est interdit, c’est le point de vue unique, lequel éclate au profit d’une multitude de perspectives possibles. **Autrement dit**, avec la photographie puis le cinéma se produit une rupture avec l’idéologie balzacienne dont ne s’est pas encore totalement émancipé, notons-le, le postulat durkheimien de l’objectivité. [ant-art-265-body]
- [127] La première vise la parité et la réciprocité, la seconde le toujours-plus. Mais tandis que la logique de « reproduction » suppose la séparation entre kastam et bisnis, celle de « développement » induit la conjonction des deux domaines (Foster, 1995 : 228). Mieux, l’argent gagné dans le bisnis permet de réaliser des fêtes mortuaires fastueuses et donc de vivre la kastam, en donnant du « grain à moudre » aux big men qui sont les principaux artisans de ces fêtes ; **autrement dit**, l’économie exogène permet de conforter l’ordre coutumier c’est là un des tours de passe-passe avec lesquels les ethnologues sont aujourd’hui familiarisés, ici ou là. [ant-art-57-body]

⁴²*Autrement dit* est composé d’un adverbe et d’un participe passé.

En outre, *autrement dit* accepte une portée locale (**P_L**), et dans ce cas il porte sur un syntagme nominal [129], adjectival [128] verbal ou prépositionnel (**q** est précédé de la même préposition que **p**) [130]. Il existe donc une équivalence syntaxique entre les séquences **p** et **q**.

- [128] L'identification biométrique est un objet bien curieux, à bien des égards irreprésentable, et qui n'offre que peu de prises à la pensée. C'est peut-être ce qui explique l'atonie qui l'entoure (Craipeau, Dubey, Guchet, 2004). La plupart des discours qui la concernent (en général pour en faire la promotion) commencent par la replacer au sein d'un continuum historique ou d'une histoire naturelle des techniques : la reconnaissance biométrique, « ça a toujours existé », depuis que l'on sait identifier un visage d'un simple coup d'œil. La technique est ainsi d'emblée naturalisée, **autrement dit légitimée** sans autre forme de procès. [soc-art-162-introduction]
- [129] La méthode développée par James Stimson consiste en plusieurs étapes. D'abord, il s'agit de simuler une reconstitution des séries, **autrement dit de remplir les trous dans les données**. Il existe de nombreuses méthodes de remplacement des données manquantes, mais compte tenu de la structure des données ici utilisées, ces méthodes s'avèrent insuffisantes. [spo-art-468-body]
- [130] Cette partie cherche donc à explorer les connexions entre la culture juridique et l'idéologie politique ou, **autrement dit, entre le droit et le politique**, et à partir de là, à montrer comment la sociologie du droit, aussi bien que la critique juridique, se structurent et prennent des formes différentes dans les traditions juridiques américaine et française. [soc-art-228-body]

c. Position syntaxique

Autrement dit se trouve uniquement en position initiale, qu'il porte sur une proposition ou sur un syntagme [133]. Comme nous venons de l'analyser, ce marqueur peut porter sur une phrase entière. Dans ce dernier cas, à la différence de *c'est-à-dire* et *à savoir*, *autrement dit* se situe de préférence en tête de phrase (149 occurrences sur le total de 197 occurrences de *autrement dit*) [131][132].

- [131] Le droit de vote de ces militants est central dans le dispositif : l'objectif est moins d'augmenter les capacités mobilisatrices du PS que d'organiser la future primaire présidentielle. **Autrement dit, la réforme vise moins à faciliter le pluralisme organisationnel propre au parti d'Epinay qu'à favoriser le pluralisme présidentialisé de la compétition interne au PS**. Il s'agit donc d'élargir la taille de l'organisation à partir d'une offre entièrement centrée sur le processus électoral. [spo-art-340-body]
- [132] Par ailleurs, la fractalité enseigne que plus une mesure est précise, plus cette mesure risque de diverger très rapidement. Alors que, en règle générale, c'est l'effet contraire qui est recherché. **Autrement dit, plus la résolution est élevée, plus la gamme d'échelles non fractale sera courte**. [geo-art-372-body]
- [133] Ce sont là, en effet, les seules « places des grandes affaires » et, pour équilibrer le tout, il recommande un « commerce d'intercourse », **autrement dit le commerce d'Inde en Inde, permis par la navigation à voile**, seul moyen efficace et commode dans ces contrées, à la condition que les tarifs protecteurs élevés dès les lendemains du 1er Empire soient abaissés. [his-art-341-body]

Quand *autrement dit* est en tête de phrase, il est souvent suivi d'une marque graphique comme la virgule, et en est précédé quand il se trouve au milieu de la phrase.

d. Valeurs sémantiques

D1. Valeur pragmatique

Par son sémantisme, *autrement dit* se distingue des marqueurs abordés jusqu'ici. Tandis que *c'est-à-dire* vise à établir une équivalence avec le référent de **p** et que *à savoir* le spécifie, *autrement dit* souligne une autre manière d'exprimer ce qui a été dit en **p**. Au plan discursif, ce MD donne une « seconde chance » au locuteur d'expliquer ce qu'il veut dire. Nous appelons cette valeur la valeur pragmatique de *autrement dit*.

Dans les exemples [134][135][136], la séquence **q** sert à généraliser le contenu sémantique de la séquence **p**. Par démarche de généralisation, nous entendons tous les processus qui suivent un mouvement du plus précis au plus général, ou bien d'une formulation spécifique à une formulation plus courante. Le locuteur peut recourir au marqueur *autrement dit* quand il souhaite réunir sous un concept unique tous les propriétés et caractères communs de différents objets. C'est pourquoi ce procédé de généralisation peut être rapproché de celui de dénomination métalinguistique. Dans l'exemple [136], **q** (dans *autrement dit q*) sert à caractériser les deux composants de la séquence **p**, à savoir *les milieux de toundra et taïga des régions boréales canadiennes et russes*, et *des espaces désertiques australiens*. Cette démarche de généralisation semble beaucoup plus claire dans l'exemple [135], où la séquence **q** synthétise différentes remarques évoquées précédemment dans la séquence **p**. Dans l'exemple [134], l'auteur renvoie à un rapport plus générique entre *offre* et *demande*, en s'appuyant sur l'opposition entre *consommateurs* et *producteurs*.

[134] Le consommateur apparaît en fait comme un chaînon intermédiaire dans le processus à enclencher. La démarche est en effet indirecte, intermédiée plus exactement, et transitive : il s'agit de faire changer les consommateurs pour faire changer les producteurs. **Autrement dit**, il s'agit d'essayer d'influencer l'offre par l'intermédiaire de la demande, d'arriver à faire apparaître les « signaux du marché ». [spo-art-642-body]

[135] On constate certes des variations, dues essentiellement aux effectifs constituant les cohortes, mais globalement, les mouvements suivent de près ceux constatés pour l'indice global de tolérance. Ainsi, toutes les cohortes semblent évoluer en conséquence en 2005 (entre -3 et -6 points), mais elles connaissent toutes depuis une augmentation régulière de la tolérance (évoluant entre +9 et +12 points). **Autrement dit**, les évolutions qu'on constate au niveau global semblent bien toucher l'ensemble des strates de la population. [spo-art-468-body]

- [136] Or ces open *shrublands* correspondent en réalité, pour une large part, soit aux milieux de toundra et de taïga des régions boréales canadiennes et russes, soit aux espaces désertiques australiens, **autrement dit** à des régions qui, tant par leur éloignement que par leurs conditions écologiques, ne présentent guère de potentialités de boisement. Les principales régions offrant en définitive des potentiels se situent dans les espaces semi-arides de Chine et des États-Unis et pour partie d’Australie. [geo-art-110-body]

Nous constatons que dans ces derniers exemples, la taille de la portée [134] est supérieure à celle du contexte gauche [134], ce qui est l’inverse dans les exemples [137][138][139], où la portée du marqueur est inférieure à celle du contexte gauche quand *autrement dit* introduit la démarche de spécialisation. Cette démarche fonctionne dans le sens inverse de celle de généralisation, la séquence **q** – la portée de *autrement dit* – tente de spécifier le contenu sémantique de la séquence **p**. Ce marqueur joue le même rôle dans les exemples [137][138]. La reformulation établie par *autrement dit* correspond aux exemples plus concrets pour des idées plus générales dans la séquence **p**. Par exemple, dans l’exemple [137] où l’auteur insiste, dans sa reformulation, davantage sur le rapport entre *l’élément k* et *la consommation*. Grâce à la reformulation instaurée par *autrement dit*, l’auteur peut donc apporter de plus amples explications et informations aux lecteurs. Dans l’exemple [139], le rôle de *autrement dit* est assez clair. En effet, la séquence **q** consiste en un exemple concret de la règle précédemment abordée par l’auteur. Dans la séquence **q**, parmi les deux langues citées, le *kabyle* est considéré comme la langue minoritaire par rapport à l’*arabe*.

- [137] La consommation de première période de vie est une fraction constante du salaire réel : Cette consommation est une fonction croissante de *k*. **Autrement dit**, lorsque *k* augmente la consommation de première période de vie augmente tandis que la consommation de seconde période de vie diminue : comme les agents accordent relativement plus d’importance à celle-ci, leur utilité baisse. [eco-art-223-body]
- [138] Nous attendions un changement d’attitude dans la situation ultrason/centration. **Autrement dit** les sujets de cette situation devraient avoir une attitude plus favorable à la sélection que les sujets des deux conditions contrôles (hypothèse principale 1 et 1 bis) et que les sujets de la situation de fausse attribution standard, c’est-à-dire de la situation ultrason/sans centration (hypothèse principale 2). En d’autres termes, en présence d’une source de fausse attribution, nous nous attendions à retrouver l’effet classique de dissonance chez les sujets qui ont été recentrés sur leur comportement contrattitudinel. [psy-art-428-body]
- [139] Dans le cas où deux langues étaient mentionnées, le glossaire de l’enquête précisait laquelle choisir : « si français + une autre langue, mettre le français en second » (donc attribuer à l’individu une appartenance étrangère puisque seule la langue citée en premier servait au classement comme indiqué plus haut) ; « lorsque deux langues maternelles étaient citées, nous avons privilégié la langue minoritaire » (**autrement dit** le kabyle si l’arabe et le kabyle avaient été mentionnés). [ant-art-312-body]

Quand *autrement dit* a la valeur pragmatique, il a une portée globale et son contexte gauche est également de nature propositionnelle. Nous observons donc que *autrement dit* peut

introduire deux démarches inversées, l'une de généralisation [134][135][136] et l'autre, de spécialisation du contenu sémantique de la séquence **p** [137][138][139].

Cette valeur pragmatique est spécifique à *autrement dit*. Dans les exemples ci-dessus, *autrement dit* ne peut pas être remplacé par *à savoir*, car l'auteur envisage de spécifier le contenu sémantique de la séquence **q**. Quant à *c'est-à-dire*, il peut remplacer *autrement dit* uniquement dans l'exemple [136] où *autrement dit* porte sur un syntagme prépositionnel. Dans les autres exemples, l'utilisation de *c'est-à-dire* semble moins naturelle. Au plan sémantique, *c'est-à-dire* ne peut pas exprimer clairement une démarche de généralisation ni de spécification comme *autrement dit*.

D2. Valeur métalinguistique

Autrement dit peut également avoir une valeur métalinguistique, mais uniquement quand ce marqueur a une portée locale. La reformulation (dans *autrement dit q*) est effectué par le rapport de synonymie [141] ou de définition [140][142].

- [140] Une dimension physique : dès le XVII^e siècle le terme de réseau est utilisé pour caractériser une structure textile croisée, « un tissu à mailles très larges », **autrement dit un entrelacement de lignes, de bandes, de fils**. Plus généralement le réseau peut être envisagé comme un assemblage de petites unités dispersées, reliées entre elles de façon plus ou moins formelle qui évoque l'image de ramification. [eco-art-254-body]
- [141] Le coût organisationnel de l'engagement **autrement dit, la contrepartie statutaire à l'adhésion** est par là des plus réduits pour le PS : une partie seulement des membres de la communauté élargie créée sur la Coopol en disposent. Dans cette perspective, communauté partisane physique et communauté partisane virtuelle ne se confondent pas, bien que la seconde englobe la première. [spo-art-340-body]
- [142] Au Canada comme en Grande-Bretagne, depuis la promulgation du projet de loi C-6, la GPA doit rester totalement altruiste, **autrement dit sans compensation financière**, alors qu'auparavant une somme forfaitaire pouvait être donnée à la mère de substitution. [ant-art-352-body]

Dans ces exemples, il s'agit d'une valeur sémantique partagée par les trois marqueurs analysés jusqu'ici. Pourtant, l'emploi de *à savoir* ne peut pas fonctionner ici, et seul *c'est-à-dire* peut remplacer *autrement dit*. Ceci correspond à la valeur métalinguistique qui introduit la définition que nous avons mise en évidence pour *c'est-à-dire*.

D3. Valeur interprétative

Autrement dit peut introduire une séquence à valeur sémantique interprétative, où la séquence **q** (*autrement dit q*) correspond à une interprétation de l'auteur [143][144].

- [143] L'objectif de la présente étude est double : il s'agit premièrement de voir comment se manifestent linguistiquement dans les textes les énoncés définitoires (**autrement dit** quelles sont les structures linguistiques qui expriment ces énoncés) (§ 2.) et deuxièmement d'envisager le problème de l'extraction automatique des segments textuels conformes à ces structures linguistiques (§ 3.). [lin-art-169-introduction]
- [144] Berlin a été choisi comme siège des institutions, mais la motion votée a permis aux partisans de Bonn, dès la deuxième moitié de l'année 1991, de tirer parti du consensus hérité du vote du 20 juin. **Autrement dit**, dès les premiers mois qui suivent le débat sur la capitale (Hauptstadtdebatte), on assiste à l'émergence d'une géographie de la pratique politique bipolaire, et finalement assez respectueuse de la tradition allemande du polycentrisme. [geo-art-550-body]

Dans les exemples [143][144], il n'existe pas de démarche de généralisation, ni de spécification, mais un rapport d'équivalence. L'auteur cherche en effet simplement d'autres moyens linguistiques équivalents pour clarifier la séquence **p**. Dans ces deux exemples, l'emploi de *c'est-à-dire* et *à savoir* peut être accepté.

Pour finir, nous relevons quelques régularités entre la portée de *autrement dit*, son contexte gauche et ses valeurs sémantiques. Si la séquence **p** et **q** a une portée locale et quasiment la même taille, *autrement dit* peut avoir une valeur métalinguistique ou interprétative. La reformulation établie par *autrement dit* peut correspondre à une définition [145] ou à une valeur interprétative [146][147].

- [145] Plusieurs bénéfices interdépendants sont espérés : soutien mémoriel, restitution de la situation (ou contexte) configurée dans l'action par les observés, ou **autrement dit** la dynamique de l'activité, réflexivité accrue des personnes interviewées, co-analyse participant(s)/chercheur/se. L'autoconfrontation croisée amplifierait ces deux dernières qualités. Plus encore, en milieu professionnel, lors des échanges entre pairs, la révélation d'écarts par rapport aux prescriptions, des pratiques implicites, est attendue (Theureau et Jeffroy, *ibid.* ; Clot, *ibid.* ; Leplat, 2000). [infcom-art-473-body]
- [146] Dans un marché commun, la politique économique se doit de privilégier l'équilibre de la balance des paiements, la lutte contre l'inflation et la stabilité des taux de change. **Autrement dit**, avec l'avènement de la Communauté économique européenne, c'est le principe même de la planification à moyen et à long terme qui se trouve sapé dans ses fondements [86]. Voilà déjà une première victoire pour les tenants du marché libre. [spo-art-637-conclusion]
- [147] Dans une situation inspirée du paradigme du garçon de café où les enfants déposent un objet tenu dans la main gauche, sur un plateau posé sur leur main droite, nous avons également observé que le lestage est précédé d'une flexion du coude droit dès 20 mois (Jover, 2002). Nous avons interprété cette flexion comme un APA, **autrement dit** comme le déploiement d'une force de direction inverse à celle résultant du lestage et permettant d'en limiter les répercussions. [psy-art-579-body]

Si la portée de la séquence **q** est supérieure à celle de **p**, *autrement dit* introduit une démarche de généralisation [148]. Au contraire, si la portée de **q** est inférieure à celle de **p**, *autrement dit* exprime une démarche de spécialisation [149].

- [148] Finalement, il s'agit ainsi de passer d'une politique de promotion des exportations, basée sur l'exportation de produits agricoles, de matières premières ou de produits manufacturés à faible contenu technologique, à une politique de substitution des exportations, qui consiste à exporter des produits manufacturés à fort contenu technologique. **Autrement dit**, l'objectif pour les économies latino-américaines est d'acquérir une compétitivité authentique, systémique non plus fondée sur la « rente périssable » constituée par les ressources naturelles et les bas salaires, mais sur la « rente dynamique » issue de l'incorporation du progrès technique dans l'activité productive. [eco-art-275-body]
- [149] Il me semble que reconnaître cette dépendance de l'interprétation par rapport à des typifications psychologiques, d'une part, et l'existence d'un savoir social, le savoir nomologique d'expérience d'autre part, c'est offrir au sociologue ou au lecteur le pouvoir de critiquer l'attribution. **Autrement dit**, analyser les capacités, compétences, émotions dont a été dotée la marionnette qui sert de modèle pour comprendre la réalité, voilà qui permet, à mon sens, de contrôler l'interprétation et d'affiner la compréhension explicative. [soc-art-259-body]

Le Tableau 5.3 : Récapitulatif des propriétés syntaxiques et sémantiques *de autrement dit* systématise les propriétés syntaxiques et sémantiques de *autrement dit*.

Paramètres	Propriétés
Position	Initiale. Au niveau phrastique, il peut se trouver en tête de phrase ou en position postverbale.
Portée	<p>Globale. Quand q porte sur une proposition, p est aussi de nature propositionnelle.</p> <p>Locale. Quand q porte sur un syntagme, p correspond également à un syntagme.</p> <div style="text-align: center;"> $\left(\begin{array}{c} \text{proposition} \\ \mathbf{p} \end{array} \right) \quad \text{Autrement dit} \quad \left(\begin{array}{c} \text{proposition} \\ \mathbf{q} \end{array} \right)$ $\left(\begin{array}{c} \left(\begin{array}{c} \text{proposition} \\ \mathbf{p} \end{array} \right) \\ \text{autrement dit} \end{array} \right) \quad \left(\begin{array}{c} \text{proposition} \\ \mathbf{q} \end{array} \right) \text{ phrase}$ $\left(\begin{array}{c} \left(\begin{array}{c} \text{Syntagme nominal} \\ \text{Syntagme verbal} \\ \text{Syntagme prépositionnel} \\ \text{Syntagme adjectival} \end{array} \right) \\ \text{autrement dit} \end{array} \right) \quad \left(\begin{array}{c} \left(\begin{array}{c} \text{Syntagme nominal} \\ \text{Syntagme verbal} \\ \text{Syntagme prépositionnel} \\ \text{Syntagme adjectival} \end{array} \right) \\ \mathbf{q} \end{array} \right) \text{ phrase}$ </div>
Valeurs sémantiques	<p>✓ <i>Valeur pragmatique</i> : uniquement quand <i>autrement dit</i> a une portée globale, <i>autrement dit</i> peut établir deux démarches inverses (généralisation ou spécialisation).</p> <p>❖ <u>Généralisation</u> : la reformulation correspond à une généralisation du contenu sémantique de la séquence p.</p> <p>[proposition]_p Autrement dit_{V.PRAGMATIQUE} [proposition]_q</p> <p><p><<q></p> <p>❖ <u>Spécialisation</u> : <i>autrement dit</i> relie quelque chose de plus général à quelque chose de plus spécialisation.</p> <p>[proposition]_p Autrement dit_{V.PRAGMATIQUE} [proposition]_q</p> <p><p><<q></p>

	<p>✓ <i>Valeur métalinguistique</i> : <i>autrement dit</i> établit des rapports de synonymie ou de définition.</p> <p>[SN/SV/SA/SP]_p Autrement dit_{V.METALINGUISTIQUE} [SN/SV/SA/SP]_q</p> <p>Portée locale</p> <p><p> = <q></p>
	<p>✓ <i>Valeur interprétative</i> : la reformulation établie par <i>autrement dit</i> correspond à une interprétation du locuteur.</p> <p>[proposition]_p Autrement dit_{V.INTERPRETATIVE} [proposition]_q</p> <p>Portée globale</p> <p><p><<q></p>

Tableau 5.3 : Récapitulatif des propriétés syntaxiques et sémantiques de *autrement dit*

En bref, *autrement dit* peut posséder trois valeurs sémantiques principales à savoir la valeur pragmatique, la valeur métalinguistique et la valeur interprétative (Tableau 5.3). La valeur pragmatique consiste en la valeur sémantique spécifique de ce marqueur et permet de distinguer *autrement dit* de *c'est-à-dire* et *à savoir*.

5.3.4. En d'autres termes

a. Définition

En d'autres termes est un peu moins fréquent que les autres marqueurs dans le même groupe, puisqu'il ne totalise que 119 occurrences. La signification du marqueur est assez transparente et est étroitement liée au sémantisme de ses composants. La définition du TLFi de *en d'autres termes* indique qu'il sert à « exprimer la même idée à l'aide d'autres mots » (d'après le TLFi). En fait, cette définition ne permet pas de le distinguer de *autrement dit* et nous souhaitons mieux le distinguer par des analyses plus précises.

b. Portée syntaxique

En raison de sa structure syntaxique, *en d'autres termes* prend en général sous sa portée une proposition entière. Le contexte gauche du marqueur est pratiquement toujours de nature propositionnelle [150][151].

[150] Cet article vise à analyser la double perturbation qu'engendrent dans une famille la présence d'un adolescent handicapé mental et, sur un temps plus court, celle d'un ethnologue. Je voudrais montrer que la réflexivité de l'ethnologue est, ici plus qu'ailleurs, étroitement liée aux résultats mêmes de l'enquête. **En d'autres termes, le retour réflexif sur les conditions de réalisation de l'enquête m'en a presque autant appris sur les organisations familiales que j'analysais que le contenu des entretiens et des observations que j'ai pu mener.** J'étudierai successivement ces deux perturbations, afin de montrer combien ces deux dimensions peuvent être intimement liées dans l'analyse. [ant-art-218-body]

[151] Dans le même ordre d'idée, il ne semble exister aucune corrélation entre le coefficient familial et la fréquence du don même si son importance est inversement proportionnelle à la présence d'enfants ou de petits-enfants survivants ; **en d'autres termes, une forte descendance incite à l'économie de moyens, et non au sacrifice de la charité elle-même.** [his-art-63-body]

En d'autres termes peut également porter sur un syntagme nominal [152], mais cet emploi est marginal dans notre corpus (4 occurrences sur le total de 119). Le contexte gauche du marqueur a la même nature syntaxique que sa portée et est introduit par la même préposition que celle-ci.

[152] Les vœux de tous les hommes de bien sont pour l'éducation rationnelle, **en d'autres termes, pour l'éducation proprement dite...** [sed-art-3-body]

c. Position syntaxique

En d'autres termes se trouve uniquement en position initiale, il est toujours suivi de sa séquence reformulatrice **q**. À l'instar de *autrement dit*, *en d'autres termes* est en général en tête de phrase et suivi de la virgule [153][154]. L'utilisation de la virgule n'est pas linguistique, mais typographique. Ce signe *démarcatif*⁴³ permet de séparer des groupes de fonctions différentes (Riegel *et al.*, 2009 : 150). Nous pouvons interpréter le marqueur *en d'autres termes* en une relative explicative comme « en utilisant d'autres termes », et celle-ci ne fait pas partie de la phrase de base. Pour l'enseignement des MD, il est important d'introduire cet aspect, car les étudiants ne le savent souvent pas, surtout ceux qui viennent de traditions graphiques sans ponctuation (ou très peu).

[153] Dans le présent modèle, nous ne proposons pas une analyse de la formation endogène du chômage ; nous nous intéressons à l'interaction entre la localisation endogène des

⁴³ Terme utilisé par Riegel, Pellat, & Rioul (2009 : 146).

chômeurs, des employés et des entreprises. **En d'autres termes**, nous ouvrons la boîte noire qu'est le CBD et nous en analysons ce qui se passe à l'intérieur. [eco-art-272-introduction]

- [154] La délégation d'un contrat strictement anti-inflationniste à la banque centrale, n'améliore donc pas forcément la situation, puisqu'il faut comparer la perte en dépenses publiques et le gain en inflation. **En d'autres termes**, un contrat d'inflation nulle peut être contreproductif. La raison est que ce type de contrat n'est pas le meilleur pouvant être atteint en l'absence de possibilité d'engagement fiscal. Définissons comme « m-optimal » un tel contrat, qui est celui qui arbitre entre le biais moyen de dépenses publiques et le biais inflationniste. [eco-art-551-body]

d. Valeurs sémantiques

D1. Valeur pragmatique

Comme l'indique son sémantisme, *en d'autres termes* permet d'insister sur la manière d'exprimer la séquence **p** par d'autres mots. La valeur pragmatique est la valeur la plus représentative de ce marqueur. La reformulation établie par *en d'autres termes* est considérée comme une formulation plus courante, plus explicite pour une formulation plus spécialisée dans son contexte gauche **p** [155][156].

- [155] Le premier problème est de savoir où est la limite de la signification linguistique. Deux solutions sont ici envisageables : la première inclut dans la signification linguistique la valeur en usage ; la seconde limite la signification linguistique à la signification compositionnelle. **En d'autres termes**, la première solution intègre le contexte dans le calcul de signification et associe à la somme signification + contexte le sens linguistique, alors que la seconde sépare signification linguistique et contexte, et explique par inférence, à savoir à partir du contexte et de la signification linguistique, le sens de l'énoncé. [lin-art-509-body]
- [156] Le développement de la compréhension, de la pensée représentative et le rappel en mémoire marquent le début de cette quatrième étape. Le jeune enfant peut désormais modifier par lui-même son comportement en fonction de ce qui est acceptable ou non de l'adulte. **En d'autres termes**, la pensée représentative rendrait l'enfant de plus en plus capable, même en l'absence de l'adulte, de se soumettre à certaines exigences, de remettre à plus tard certaines activités et d'inhiber certains comportements moteurs ou verbaux, ce qui représente une étape majeure vers l'acquisition d'un système intériorisé de régulation. [psy-art-382-body]

À la différence de *autrement dit*, *en d'autres termes* ne traduit pas forcément la démarche de généralisation. *En d'autres termes* manifeste le souci du locuteur qui va à la recherche d'autres formulations, de mots nouveaux pour apporter plus de clarté à son énoncé. Dans l'exemple [155], la séquence **q** (dans *en d'autres termes q*) sert à clarifier les notions spécifiques évoquées dans la séquence **p** par une formulation plus simple. Dans l'exemple [156], *en d'autres termes* introduit les explications de l'auteur sur l'idée qui vient d'être abordée dans **p** sur la *pensée représentative*. Sur le plan syntaxique, quand *en d'autres termes*

introduit cette valeur, il a une portée globale et son contexte gauche **p** est également de nature propositionnelle.

Dans l'exemple [157], l'emploi de *en d'autres termes* sert également à apporter plus de précisions.

- [157] L'un des objectifs de notre seconde expérience sera donc de tester cette hypothèse en répliquant l'expérience 1, mais dans un contexte de comparaison interpersonnelle. Si, comme nous venons de le suggérer, lors de comparaisons descendantes dans un contexte intergroupe, les individus se centrent sur leur identité personnelle plutôt que sur leur identité sociale, nous devrions reproduire, avec des comparaisons descendantes interpersonnelles, les résultats observés avec des comparaisons descendantes dans un contexte intergroupe (expérience 1). **En d'autres termes, dans un contexte de comparaisons descendantes interpersonnelles, les individus devraient présenter le même niveau d'estime de soi que leur cible de comparaison soit de leur groupe ou d'un autre groupe.** [psy-art-563-body]

La reformulation établie par *en d'autres termes* correspond plutôt à une démarche de spécialisation qu'à une généralisation. Dans ces exemples, seul le marqueur *autrement dit* peut remplacer *en d'autres termes*. L'emploi de *c'est-à-dire* et *à savoir* n'est pas possible.

D2. Valeur résumante

Pour *en d'autres termes*, nous relevons une valeur sémantique moins présente pour les autres MD, la valeur résumante [158][159].

- [158] Or la plupart des définitions du paysage existantes s'appuient sur le caractère visuel de la notion, mettant en rapport un sujet observant et un objet observé (Rougerie, Berouchachvili, 1992). En ce sens, le paysage ne correspond pas à un espace géométrique « classique », à une surface homogène. Tout point n'est pas équivalent en matière paysagère, en raison à la fois de sa visibilité physique à partir d'autres lieux et de l'existence de parcours privilégiés, de points de vue, permettant la perception dudit paysage à un nombre élevé d'observateurs. **En d'autres termes, le paysage s'entend comme un réseau constitué de points de vue, d'itinéraires et d'angles morts de l'observateur.** [geo-art-393-body]
- [159] Au sein de ces débats, revient ainsi de manière récurrente une interrogation sur l'« introduction des nouvelles technologies » dans l'entreprise à laquelle se trouvent immédiatement liées la formation, la motivation et surtout l'information et la consultation des travailleurs. **En d'autres termes, l'introduction des nouvelles technologies n'est pas conçue comme une pure adaptation des entreprises européennes à un modèle s'imposant de l'extérieur.** Certes, le modèle japonais revient de manière fréquente. Cependant, les travailleurs sont envisagés comme des agents actifs pour la réussite de ces mutations technologiques. [soc-art-405-body]

Dans l'exemple [158], la séquence **p** dépasse le cadre d'une seule phrase, elle se compose de plusieurs phrases *Or la plupart des définitions... à un nombre élevé*

d'observateurs. La séquence reformulatrice **q** établie par *en d'autres termes* a une portée globale et joue le rôle de résumé de ce qui a été énoncé dans la séquence **p**. Le marqueur *en d'autres termes* dans [159] joue un peu le même rôle que dans [158]. L'auteur cherche à résumer ce qui a été dit pour ensuite développer ses arguments. Dans ces deux exemples, *en d'autres termes* ne peut être remplacé par les trois marqueurs précédents. À *savoir* ne fonctionne pas, car l'auteur n'envisage pas de spécifier la séquence **q**. *C'est-à-dire* peut introduire une valeur conclusive, mais l'utilisation de *c'est-à-dire* ne peut pas fonctionner, car ce marqueur ne permet pas d'introduire les arguments de l'auteur. L'usage de *autrement dit* dans ces exemples est moins naturel, et ce marqueur ne permet pas de véhiculer l'idée de résumé qu'*en d'autres termes* cherche à dégager.

D3. Valeur métalinguistique

En d'autres termes peut avoir une valeur métalinguistique de synonymie. Notons que cette valeur est beaucoup moins fréquente que les autres valeurs abordées pour *en d'autres termes* [160][161].

[160] Si les actionnistes étaient cohérents avec eux-mêmes, suggère Sperber, ils adopteraient le programme naturaliste et mécaniste. Sur ce point, le raisonnement de l'anthropologue est peut-être un peu radical. Il fait penser à ce parallogisme que les scolastiques nommaient *a dicto secundum quid ad dictum simpliciter* qui met en garde contre les conclusions abusives d'une proposition prise isolément, **en d'autres termes**, contre la négligence de la conditionnalité d'une des prémisses du raisonnement. [soc-art-60-body]

[161] Je vais maintenant tenter de « dégager une image la plus complète possible » de la discipline, **en d'autres termes** de décrire ce qu'est l'anthropologie en France aujourd'hui. Il serait utopique de prétendre accomplir ce programme tant sont nombreuses et variées les productions anthropologiques et les opinions émises par les uns et les autres sur leur nature et leur direction. Je me contenterai donc de quelques remarques sur des points non dénués d'importance. [ant-art-133-body]

En d'autres termes introduit cette valeur uniquement quand il a une portée locale. Dans les exemples [160][161], ce marqueur porte sur un syntagme prépositionnel, et les séquences **p** et **q** sont précédées de la même préposition. Dans ces exemples, *en d'autres termes* peut être remplacé par les trois autres marqueurs *autrement dit*, *c'est-à-dire*, *à savoir*.

En d'autres termes prend des propositions sous sa portée dans la majorité des cas. En plus, comme les séquences **p** et **q** (dans la suite **p en d'autres termes q**) peuvent se composer d'une ou plusieurs propositions [158], il nous est difficile de conclure sur la régularité des valeurs sémantiques de ce marqueur en relation avec la taille de la séquence **p** et **q**. En dehors

du cas où *en d'autres termes* introduit une séquence à valeur métalinguistique, la taille de **p** et celle de **q** sont égales.

Paramètres	Propriétés
Position	Initiale. Au niveau phrastique, <i>en d'autres termes</i> est placé en tête de phrase ou en position postverbale.
Portée	<p>Globale. Il porte essentiellement sur une proposition, la séquence p est aussi de nature propositionnelle.</p> <p>Locale (très peu). <i>En d'autres termes</i> porte sur un syntagme nominal et la séquence p l'est aussi.</p> <div style="text-align: center;"> $\left(\begin{array}{c} \text{proposition} \\ \text{p} \end{array} \right) \text{ En d'autres termes } \left(\begin{array}{c} \text{proposition} \\ \text{q} \end{array} \right)$ $\left(\left(\begin{array}{c} \text{proposition} \\ \text{p} \end{array} \right) \text{ en d'autres termes } \left(\begin{array}{c} \text{proposition} \\ \text{q} \end{array} \right) \right)_{\text{phrase}}$ $\left(\left(\begin{array}{c} \text{Syntagme} \\ \text{nominal} \\ \text{p} \end{array} \right) \text{ en d'autres termes } \left(\begin{array}{c} \text{Syntagme} \\ \text{nominal} \\ \text{q} \end{array} \right) \right)_{\text{phrase}}$ </div>
Valeurs sémantiques	<p>✓ <i>Valeur pragmatique</i> : <i>en d'autres termes</i> met le lien entre une formulation plus spécifique et une formulation plus générale. Il a une portée globale.</p> <p>[proposition]_p En d'autres termes_{V.PRAGMATIQUE} [proposition]_q</p> <p>✓ <i>Valeur résumante</i> : la séquence reformulatrice q joue le rôle de résumé. <i>En d'autres termes</i> porte sur une proposition.</p> <p>[proposition₁+proposition₂+...] _p En d'autres termes_{V.RESUMANTE} [proposition]_q</p> <p>✓ <i>Valeur métalinguistique</i> : <i>En d'autres termes</i> établit un rapport de synonymie entre la séquence p et q. Il a une portée locale, il porte sur un syntagme prépositionnel.</p>

	[SP] _p en d'autres termes V.METALINGUISTIQUE [SP] _q
--	---

Tableau 5.4 : Récapitulatif des propriétés syntaxiques et sémantiques de *en d'autres termes*

En bref, le marqueur *en d'autres termes* a principalement une valeur pragmatique. Ce marqueur traduit l'intention du locuteur d'apporter plus de clarté ou plus de précision à son discours, dans une formulation plus générale ou plus claire selon le contexte. Le locuteur ne cherche pas à donner des informations nouvelles au lecteur, mais l'utilisation de nouveaux mots permet de clarifier le discours. Outre cela, il peut également avoir une valeur conclusive et une valeur métalinguistique [160]. En général, *en d'autres termes* est susceptible d'être remplacé par *autrement dit*, mais il peut uniquement être remplacé par *c'est-à-dire* et *à savoir* quand il introduit une séquence à valeur métalinguistique. Une autre propriété de *en d'autres termes* qui permet de le distinguer de *autrement dit* est la nature syntaxique du contexte gauche de chaque marqueur. La séquence **p** de *en d'autres termes* peut être composée de plusieurs phrases, tandis que celui de *autrement dit* se limite souvent à une phrase.

Remarques conclusives

Nous pouvons d'abord affirmer que les MD *c'est-à-dire*, *à savoir*, *autrement dit* et *en d'autres termes* sont interchangeables quand la reformulation établie par ceux-ci possède une valeur métalinguistique et est instaurée par un rapport de synonymie. Ensuite, tandis que *à savoir* apporte de nouvelles informations, de nouvelles connaissances dans la séquence reformulatrice, *c'est-à-dire*, *autrement dit*, *en d'autres termes* peuvent uniquement reformuler tout ce qui est déjà connu du lecteur ; soit l'information a été abordée dans le texte, soit ces informations sont supposées connues du lecteur.

Les propriétés saillantes de chaque marqueur dans cette catégorie sont représentées dans le Tableau 5.5 : Comparaison des caractéristiques de *c'est-à-dire*, *à savoir*, *autrement dit* et *en d'autres termes*.

MD	Propriétés syntaxiques et sémantiques
C'est-à-dire	- <i>C'est-à-dire</i> introduit un contenu supposé comme connu par le lecteur, ce

	<p>qui se manifeste soit par l'utilisation du déterminant défini dans la portée et dans le contexte gauche, soit par l'emploi du déterminant indéfini <i>un</i> à valeur représentative, soit par l'apparition des renvois à des parties dans le texte, etc.</p> <p>- <i>C'est-à-dire</i> fonctionne comme un MD générique, ce qui explique son omniprésence dans le corpus. C'est le marqueur le plus productif dans cette catégorie et il peut exprimer de différentes valeurs sémantiques.</p>
À savoir	<p>- La <i>valeur énumérative</i> est la valeur principale de <i>à savoir</i>. La séquence reformulatrice q correspond à une énumération. Le marqueur est souvent précédé d'adjectifs cardinaux.</p> <p>➤ La reformulation introduite par <i>à savoir</i> apporte quelque chose de nouveau.</p> <p>➤ Il peut être remplacé par <i>c'est-à-dire</i> uniquement quand la reformulation est supposée comme connue ou abordée dans le texte par l'utilisation d'adjectifs comme <i>tel</i>, du déterminant défini, ou du pronom démonstratif, etc.</p> <p>- <i>À savoir</i> peut être remplacé par <i>c'est-à-dire</i> quand il établit un rapport de synonymie entre p et q.</p>
Autrement dit	<p>- Ce marqueur est orienté vers le lecteur et la <i>valeur pragmatique</i> est la valeur la plus courante. Il porte dans ce cas sur une proposition.</p> <p>➤ Le rapport entre la séquence p et q est instauré dans une démarche de généralisation et de spécialisation.</p> <p>➤ <i>Autrement dit</i> ne peut pas être remplacé par <i>c'est-à-dire</i>, <i>à savoir</i>, <i>en d'autres termes</i>, mais il peut être remplacé par <i>en d'autres termes</i> quand <i>autrement dit</i> introduit une démarche de spécialisation.</p> <p>- <i>Autrement dit</i> introduit une séquence à <i>valeur métalinguistique</i> et à <i>valeur interprétative</i>, uniquement quand il a une portée locale.</p> <p>➤ Il peut être remplacé par <i>c'est-à-dire</i>, mais pas par <i>à savoir</i> ou <i>en d'autres termes</i>.</p>

	<p>➤ À l’instar de <i>c’est-à-dire</i>, il permet de reformuler ce qui a été dit ou connu.</p>
En d’autres termes	<p>- <i>En d’autres termes</i> manifeste l’intention du locuteur qui est à la recherche d’une formulation plus compréhensible. Il a une portée globale.</p> <p>➤ Dans ce cas, <i>en d’autres termes</i> peut être remplacé par <i>autrement dit</i>, mais pas par <i>c’est-à-dire</i>, à savoir.</p> <p>- Quand <i>en d’autres termes</i> a une <i>valeur résumante</i>, il porte sur une proposition. Son contexte gauche p peut se composer de plusieurs phrases.</p> <p>➤ Ce marqueur ne peut pas être remplacé par d’autres marqueurs de reformulation quand il introduit une séquence reformulatrice q ayant cette valeur.</p> <p>- Peu d’occurrences de <i>en d’autres termes</i> ont une <i>valeur métalinguistique</i>, mais <i>en d’autres termes</i> a une portée locale dans ce cas.</p> <p>➤ Il peut être remplacé par <i>c’est-à-dire</i>.</p>

Tableau 5.5 : Comparaison des caractéristiques de *c’est-à-dire*, à savoir, *autrement dit* et *en d’autres termes*

5.4. Récapituler

Les trois MD de ce groupe (*en somme*, *en résumé*, *en un mot*) ont comme rôle commun de résumer les idées développées antérieurement et d’indiquer les points essentiels dans l’objectif de faciliter la compréhension du contenu pour le lecteur. L’ordre d’analyse présenté ici s’appuie sur leur fréquence, en commençant par *en somme* qui reste le marqueur le plus fréquent. Les autres marqueurs vont être comparés avec celui-ci afin de mieux faire ressortir le fonctionnement syntactico-sémantique de chaque MD.

5.4.1. En somme

a. Définition

En somme est le marqueur le plus fréquent et le plus représentatif de cette catégorie (57 occurrences). Dans le TLFi, *en somme* est défini comme un synonyme de *finalement* et de *tout compte fait*. La signification du marqueur est liée au sémantisme de son composant

somme, terme d'abord utilisé en mathématiques pour indiquer le résultat d'une addition. Pourtant, *en somme* ne peut être considéré comme une simple addition de différents points de vue ou idées ou « le point focal de cette série de points de vue », comme ce qu'affirme Rossari (Rossari, 1994 : 50). La séquence **q** introduite par *en somme* consiste selon nous en une synthèse des idées que l'auteur a précédemment abordées.

b. Portée syntaxique

En raison de son sémantisme, l'étendue de la séquence **p** de *en somme* est plus ou moins large et de nature propositionnelle (mais il peut s'agir aussi d'un paragraphe entier ou même de plusieurs paragraphes) tandis que la séquence **q** prend en général sous sa portée une proposition (**P_G**) [162][163].

[162] Force est de constater que les jeux didactiques font partie de la réalité quotidienne des pratiques physiques scolaires tant du point de vue de l'analyse des pratiques enseignantes observables (cf. tableau 2) que du point de vue de la pertinence et de la faisabilité didactique. En somme, les pratiques physiques sociales sont transformées en pratiques scolaires au gré des impératifs éducatifs, des contraintes pédagogiques et de la créativité de l'enseignant. [sed-art-52-body]

[163] Examinons à présent le couple pureté/impureté dans le cas des Hébreux ordinaires. On constate immédiatement que les occasions de se rendre impur sont innombrables, mais que la logique qualitative binaire présence/absence s'accompagne d'évaluations quantitatives: les infractions aux lois de pureté sont d'une gravité inégale, que traduit la gradation de leurs conséquences. Tout se passe en somme comme si chaque loi instituait son registre de pureté/impureté, l'effet des transgressions étant la perte d'une sorte seulement de pureté alimentaire, sexuelle, etc. [ant-art-368-body]

Nous relevons également quelques occurrences où *en somme* porte sur un syntagme, par exemple dans l'exemple [164] où il porte sur un syntagme verbal. Dans ce cas, *en somme* a une portée locale (**P_L**) et la taille de son contexte gauche est plus restreinte et se limite au niveau d'un syntagme.

[164] Ici encore, les emprunts, les imitations exploitent un code préexistant pour lui donner de nouvelles dimensions à mesure qu'il est l'objet d'appropriations sociales et culturelles nouvelles. Ainsi, au sein des catégories populaires, ces multiples jeux de l'apparence, de couleurs et de parfums, témoignent d'une indéniable tendance ludique, d'une volonté d'érotiser les corps, du désir de prendre soin de soi et de se plaire, en somme de jouir de la vie. [his-art-598-body]

c. Position syntaxique

À l'instar d'autres marqueurs, *en somme* se trouve dans la plupart des cas en position initiale. En prenant sous sa portée une séquence de nature propositionnelle, il se trouve en

général en tête de phrase [165], ou en position postverbale comme dans [166]. Nous relevons une seule occurrence (sur le total de 57 occurrences) où *en somme* se situe en position finale [167], et il y a ici un changement de sens, ce MD n'introduit pas une reformulation mais un point de vue de l'auteur et peut être remplacé par *en quelque sorte*.

- [165] Cette théorie s'accorde particulièrement aux travaux menés par Luc Boltanski, Laurent Thévenot et Eve Chiapello dans le cadre des économies de la grandeur (Boltanski, Thévenot, 1991 ; Boltanski, Chiapello, 1999). Il s'agit effectivement d'une perspective voisine de ce que nous avons soulevé à propos des cadres goffmaniens puisque ces auteurs expliquent comment qualifier et évaluer les choses et les êtres en vue d'agir conjointement. Ils étudient les situations de rupture d'ordre afin de déterminer comment les acteurs font pour le rétablir. En somme, comment, par la justification, créer de la coordination, du consensus, en faisant appel pour cela à « un principe supérieur commun ». [infcom-art-581-body]
- [166] Une perspective du Nouveau Monde pourra certainement aider à aller dans ce sens ; une révision des « alliances scientifiques et techniques » est à souhaiter. En se tournant vers l'exportation, les producteurs américains, australiens et sud-africains poussent en somme à l'affirmation de nouvelles disciplines et assurent leur promotion dans les organismes internationaux. [spo-art-51-body]
- [167] Ces deux dernières années, les prix de l'immobilier ont été multipliés par cinq, sous l'influence conjuguée des transactions que nous venons de signaler, et de l'arrivée d'investisseurs britanniques, qui créent des petits ghettos de réhabilitations, dans quelques quartiers de Sofia, ou encore dans des villages ruraux, déstructurant les modes d'exploitation agricole antérieurs. Les banques bulgares, actuellement, ne prêtent pas pour de tels achats. Des petits commerces de proximité, des cafés, des coiffeurs, etc., sans lien avec les précédents « toucrédit », s'installent dans les rez-de-chaussée, et tout ce monde-là attend sa promotion au statut de « nouvelle classe moyenne », adoptant par avance les comportements supposés de celle-ci. Mais cette population est terriblement fragilisée par la stagnation de son pouvoir d'achat face à l'augmentation du foncier : la plupart de ces habitants ont contracté des emprunts auprès de leurs interlocuteurs immobiliers mafieux, avec des clauses de réajustement de la dette en fonction de l'évolution du coût de la pierre... Une « gentrification » nourrit l'autre, en somme. [soc-art-458-body]

En nous basant sur la portée et la position de *en somme*, nous relevons plusieurs valeurs sémantiques de ce marqueur.

d. Valeurs sémantiques

En somme envisage d'instaurer dans la séquence reformulatrice les éléments ou les idées communs des éléments antérieurement abordés dans son contexte gauche par deux démarches 1/ par synthèse ou par généralisation ou 2/ par caractérisation. Nous distinguons donc deux valeurs principales que peut effectuer *en somme* dont la valeur récapitulative et la valeur épistémique.

D1. Valeur récapitulative

La séquence **q** (*en somme q*) correspond à une formulation synthétique des segments antérieurs [168][169].

[168] Delphine restera quatre ans et demi dans cette entreprise, d'abord comme ingénieur de recherche puis comme « biochemist » au sein d'une antenne nouvellement créée aux États-Unis, avant de devenir « chargée de recherche prospective » auprès d'une marque de luxe concurrente. L'accès de Delphine au poste d'ingénieur de recherche succède donc à une courte période de questionnements suscités, plutôt que par l'absence de perspectives objectives (cas des étudiants les moins dotés), par le caractère très prévisible des débouchés qui lui sont offerts à l'issue d'un parcours universitaire très linéaire, sorte de voie royale (DESS obtenu cinq ans après le baccalauréat, dans la continuité des autres diplômes passés). En somme, cette courte séquence biographique de déambulation professionnelle a semble-t-il eu pour fonction de faire accepter à Delphine un destin auquel elle n'adhérerait qu'à moitié et de lui faire goûter, par contraste, les nombreux avantages de la position sociale à laquelle elle peut accéder. [soc-art-118-body]

[169] [...] Les documents manifestent soit une adéquation spontanée aux attentes implicites de l'institution (tableaux de bord, documents attestant l'existence de groupes de travail dans le REP, description d'une structure de pilotage, texte réflexif en plus des évaluations particulières de fiches-actions) ou au contraire une méconnaissance des attentes institutionnelles. Par exemple l'intégration d'éléments financiers qui n'était pas du tout attendue dans ce bilan d'étapes dont la vocation est de nature pédagogique, mais qui était souhaitée par ailleurs par le Conseil Général. La forme et l'organisation du document central et des pièces jointes attestent en somme la distance ou de la connivence avec l'institution, de l'existence ou non de canaux de circulation de l'information. [sed-art-234-body]

Dans l'exemple [168], l'introduction de *en somme* permet à l'auteur d'insister sur le rôle de la description biographique de Delphine. Dans [169], la synthèse dans la séquence reformulatrice résulte des exemples concrets sur la forme et l'organisation des documents administratifs que l'auteur a développés. Dans ces exemples, au plan syntaxique, *en somme* a une portée globale et son contexte gauche se compose de plusieurs phrases. Au plan sémantique, les éléments et idées de la séquence **p** sont concordants. À l'instar de Roulet (1987), nous constatons un changement de perspective, d'angle de vue qui fait progresser le raisonnement. Le lecteur est conduit à suivre le cheminement de la pensée du locuteur. Autrement dit, l'emploi de *en somme* devient impossible si un élément opposé s'intègre dans la séquence **p**.

Dans les exemples [170][171], l'utilisation de *en somme* donne également lieu à une synthèse qui peut servir à des objectifs distincts.

[170] Or sa présence dans le judaïsme ancien et diasporique comme dans certains aspects de la pensée chrétienne a sans doute eu plus de poids en Occident qu'à travers les constructions savantes où Descola le reconnaît théories des signatures et de la grande chaîne de l'être. Mon propos n'est donc pas du tout de critiquer le modèle lui-même ou son applicabilité aux cas du judaïsme et du christianisme, mais au contraire de montrer qu'il offre une grille de lecture d'une grande pertinence. Il s'agira en somme d'ajouter une étude de cas à celles qui ont permis à Philippe Descola de construire son

modèle : son absence du corpus de départ ayant une valeur intéressante de test si, comme je le pense, l'application de cette théorie permet de clarifier des aspects du texte biblique qui, sans cela, resteraient des énigmes. [ant-art-368-body]

- [171] [...] Enfin, peu d'élèves pratiquent le Badminton en club, ce qui équilibre les différences d'aisance motrice. Ainsi le badminton a-t-il connu le passage de la méconnaissance à la reconnaissance scolaire. Cette légitimation a d'ailleurs permis à des spécialistes du badminton d'intégrer l'université afin de former les étudiants des STAPS. **En somme**, contrairement au tennis, le badminton s'impose dans les séances d'EPS davantage pour des raisons pratiques (planification sur l'année, coût financier moindre et facilitation pour tous à entrer dans l'activité) que pour des raisons de médiatisation, de spectacle ou d'identification au champion que presque aucun élève ne connaît. [sed-art-52-body]

Dans [171], la séquence reformulatrice **q** permet à l'auteur de mieux expliquer la suprématie du badminton par rapport au tennis dans le contexte scolaire. Dans [170], l'auteur prend appui sur les arguments développés dans la séquence **p** pour souligner l'intérêt de sa méthode de mener une étude de cas. Quel que soit l'objectif de l'auteur, *en somme* a la propriété de souligner la progression des idées et la construction du raisonnement. De plus, les éléments ou idées développés dans la séquence **p** vont dans le même sens que l'idée synthétique de la séquence reformulatrice.

D2. Valeur épistémique

En somme peut avoir une valeur épistémique quand la reformulation établie par le marqueur sert à introduire une idée ou un point de vue plus général du locuteur pour ce qui a été dit dans la séquence **p** [172][173].

- [172] Certains items de notre grille de lecture sont plus différenciants que d'autres quant à la distance (au sens statistique du terme) des documents les uns par rapport aux autres. Ce sont : la présence de descriptions de groupes de travail concrets internes au REP (attestés par des descriptions précises et des documents joints) ; les autocritiques précises des pratiques pédagogiques ; la présence d'une partie rédigée en commun par les degrés d'enseignement élémentaire et secondaire. **En somme**, ce qui fait la principale différence quant à l'écriture des bilans d'étape, c'est la richesse du travail collectif réel. [sed-art-234-body]
- [173] Sont créés, autour de ces monuments et ruines, des placettes ou des jardins, pour ouvrir la vue, mettre en scène la majesté de l'espace monumental, mais aussi pour donner le sentiment du sacré, l'idée que l'on sort de l'espace quotidien et banal pour entrer dans un espace qui propose un message, qui demande à être respecté (Ricci, 2006, p. 20-21). **En somme**, l'aménagement de l'espace urbain autour du monument est ce qui entraîne un regard différent, ce qui fait la monumentalité. [geo-art-349-body]

Dans l'exemple [172], *en somme* instaure un rapport d'hyponymie entre **p** et **q**. La séquence reformulatrice correspond à une généralisation des idées développées dans la séquence **p**. Les trois items de la grille de lecture que l'auteur a évoqués renvoient tous au

travail collectif. L'emploi de *en somme* dans [173] sert aussi à généraliser l'objectif de *l'aménagement de l'espace urbain autour d'un monument*. Dans ces exemples, *en somme* met en relation des idées en allant de quelque chose de spécifique à quelque chose de général.

Dans les deux exemples suivants, *en somme* porte sur un syntagme [174] et sur une proposition [175]. Ce marqueur indique moins le mouvement de synthèse qu'une généralisation [174][175].

[174] Des poudres d'origine métallique le plus souvent plomb, mercure, étain, bismuth sont mélangées à une substance plus ou moins liquide et grasse, simple ou composée eau, huile, pommade avant d'être appliquées sur le visage. Tout autant que de modifier le coloris, il s'agit de combler les cicatrices et les rides, de lisser les aspérités, de dissimuler le rouge de l'émotion, les marques de la fatigue ou de la maladie ; de réaliser en somme de véritables masques, illisibles et impénétrables, susceptibles de servir toutes les ambitions et toutes les intrigues. [his-art-598-body]

[175] Sensibilisation, information et éducation des consommateurs s'insèrent en fait dans un raisonnement qui devient presque consensuel. Dans les propositions élaborées, cette conception est avancée de telle sorte qu'elle permet de qualifier dans le même mouvement la masse des sujets à rallier et les stratégies à engager. Ainsi intéressés et sensibilisés, les consommateurs en l'occurrence doivent pouvoir devenir en somme les agents du changement. [spo-art-642-body]

Dans la démarche de généralisation, l'auteur peut être conduit à nommer le phénomène comme dans [175]. L'emploi de *en somme* résulte de toute l'analyse du contenu dans son contexte gauche. C'est pourquoi ce marqueur est utilisé uniquement après que l'auteur aura passé en revue tous les éléments faisant l'objet de la reformulation. Dans [174], la séquence reformulatrice introduite par *en somme* englobe les idées développées dans la séquence **p**. En fait, celles-ci concourent à décrire les rôles et les bienfaits du mélange des *poudres d'origine métallique avec substance plus ou moins liquide et grasse*.

Nous constatons un trait sémantique régulier de *en somme* lié à son comportement syntaxique. Si *en somme* se trouve en position postverbale et prend sous sa portée un syntagme, il sert en général à catégoriser ou caractériser la séquence **p** [174][175][176]. Dans d'autres cas, si *en somme* porte sur une proposition, il introduit une synthèse ou une séquence à valeur récapitulative. En ce qui concerne le contexte gauche du marqueur dans les deux cas, celui de *en somme* est beaucoup plus long et peut être constitué d'un paragraphe ou même de plusieurs paragraphes quand il a une valeur récapitulative. La séquence **p** (dans **p en somme**) est plus restreinte et se limite à une ou deux phrases quand il a une valeur épistémique.

[176] Elles peuvent paraître compléter les initiatives diplomatiques, en permettant l'établissement de contacts diversifiés (par exemple au moyen du financement de la formation de leaders potentiels), en assurant une présence dans des lieux moins investis, en ajoutant à la carte « élitiste » de la diplomatie américaine des contacts plus « humbles » (dans des groupes sociaux et politiques distincts), voire une carte

« humaniste » à un jeu perçu comme guerrier et unilatéraliste ; en accroissant, en somme, l'influence de gouvernements. [spo-art-586-body]

Nous pouvons résumer les propriétés essentielles de *en somme* dans le Tableau 5.6 :
Récapitulatif des propriétés syntaxiques et sémantiques de *en somme*.

Paramètres	Propriétés
Position	Initiale pour introduire la séquence reformulatrice. Au niveau phrastique, il est souvent placé en tête de phrase et en position postverbale.
Portée	<p>Globale. Les séquences p et q sont en général de nature propositionnelle.</p> <p>Locale (très peu). <i>En somme</i> porte sur un syntagme nominal ou verbal.</p> $\left(\begin{array}{c} \text{proposition} \\ \text{paragraphe} \end{array} \right)_p \text{ En somme } \left(\begin{array}{c} \text{proposition} \end{array} \right)_q$ $\left(\begin{array}{c} \text{proposition} \\ \text{paragraphe} \end{array} \right)_p \left(\left(\begin{array}{c} \text{Début } q \end{array} \right) \text{ en somme } \left(\begin{array}{c} \text{Fin } q \end{array} \right) \right)_{\text{phrase}}$ $\left(\left(\begin{array}{c} \text{Syntagme nominal} \\ \text{Syntagme verbal} \end{array} \right)_p \text{ en somme } \left(\begin{array}{c} \text{Syntagme nominal} \\ \text{Syntagme verbal} \end{array} \right)_q \right)_{\text{phrase}}$
Valeurs sémantiques	<p>✓ <i>Valeur récapitulative</i> : <i>en somme</i> relie les idées concordantes. La reformulation correspond à la formulation synthétique. Il a une portée globale.</p> $[\text{proposition}_1 + \dots + \text{proposition}_n]_p \text{ En somme } [\text{proposition}]_q$ <p>✓ <i>Valeur épistémique</i> : <i>en somme</i> sert à généraliser ce qui a été dit dans la séquence p. Il a une portée locale.</p> $[\text{SN/SV}]_p \text{ en somme } [\text{SN/SV}]_q$

Tableau 5.6 : Récapitulatif des propriétés syntaxiques et sémantiques de *en somme*

Pour conclure, *en somme* sert à articuler des idées ou des points de vue concordants. Ce marqueur se différencie de *en conclusion*, *en dernier lieu*, etc. par les procédés de synthèse ou de généralisation qu'il opère. Si ces derniers servent soit à marquer la fin d'une étape ou soit à tirer une conclusion, la reformulation introduite par *en somme* envisage de révéler l'essentiel par une séquence plus synthétique. Ce marqueur a la particularité d'indiquer le cheminement de la pensée de l'auteur, ce qui permet de le distinguer d'autres marqueurs récapitulatifs que nous allons analyser par la suite.

5.4.2. En résumé

a. Définition

En résumé est moins fréquent que *en somme* dans notre corpus (21 occurrences au total). Cependant, il est plus utilisé dans notre corpus SHS que dans le corpus Emolex, il connaît le ratio de 0,08 dans notre corpus contre 0,02 dans le corpus Emolex. Il fonctionne comme une locution adverbiale. Sa signification est assez évidente et se décode grâce au sémantisme du substantif *résumé*. Le TLFi le définit comme un acte de résumer, ce qui nous semble un peu général et n'indique pas suffisamment ses valeurs d'usage. La séquence introduite par ce marqueur a la fonction de récapituler les éléments importants d'une séquence textuelle.

b. Portée syntaxique

En résumé a toujours une portée globale (P_G). Il ne prend que sous sa portée une proposition ou même un paragraphe. Son contexte gauche est également de nature propositionnelle, mais celui-ci peut aussi se composer d'un paragraphe entier ou de plusieurs paragraphes [177][178].

[177] Autre pression pour les étudiants des pays du Sud, celle des conditions administratives à remplir pour obtenir le renouvellement du titre de séjour chaque année ; renouvellement plus difficile en cas de redoublements répétés, ce qui oblige ces étudiants à mettre la priorité sur les études, au détriment des loisirs. **En résumé, les migrations pour études peuvent avoir des significations totalement différentes selon le continent d'origine ce qui explique pour une très grande part les différences de pratiques de mobilités entre les étudiants internationaux.** [geo-art-6-body]

[178] La plupart d'entre eux fournissent des informations sur le contexte national, les développements liés aux objectifs de Bologne et la création d'un « comité Bologne »,

relai au niveau national. Ces documents restent cependant très différents dans leur structure et hétérogènes dans le type d'informations et de données mobilisées. **En résumé, on peut dire qu'autour de 2003, les gouvernements se sentent obligés de rendre un rapport et que, dans la mesure où celui-ci est désormais librement disponible en ligne (et donc aisément accessible non seulement pour les autres États européens, mais aussi pour tous les acteurs et observateurs nationaux curieux de voir la présentation de la situation nationale faite par leur ministère), il peut devenir inconfortable pour un ministère de produire un rapport versant trop caricaturalement dans l'autopromotion.** [spo-art-117-body]

c. Position syntaxique

Toutes les occurrences de *en résumé* dans notre corpus se trouvent en position initiale et introduisent la séquence reformulatrice [179][180]. Le MD porte uniquement sur une proposition et est toujours placé en tête de phrase. Au niveau de la structure textuelle, ce marqueur est utilisé en général dans le corps du texte, et une seule occurrence de *en résumé* apparaît en introduction.

- [179] Dans La Couleur comme maléfice, Jean-Luc Bonniol cite par exemple 53 termes du lexique populaire du métissage à Haïti qui ne sont pas réductibles aux classifications pseudo-savantes utilisant la proportion de sang blanc et noir (telle celle de Moreau de Saint-Méry fondée sur les 128e parts du sang). **En résumé, aucune solution n'est satisfaisante.** [ant-art-312-body]
- [180] En d'autres termes, le nombre total de participants à une procédure d'enchères est une variable sociologique qui ne doit pas être confondue avec la taille du marché proprement dit, qui est la variable économique requise par le modèle de l'enchère. **En résumé, Baker cherche à nous convaincre que ses observations empiriques réfutent le « modèle économique idéal typique » et corroborent un modèle sociologique dans lequel un facteur social, à savoir la structuration sociale du marché, joue un rôle causal essentiel, mais négligé par les économistes.** [soc-art-552-body]

d. Valeurs sémantiques

Nous relevons dans le corpus que la valeur épistémique la plus représentative du marqueur. La reformulation établie par *en résumé* est considérée comme une formulation plus condensée des segments antérieurs. La séquence reformulatrice **q** sert à résumer tous les idées ou arguments développés et *en résumé* intervient donc au terme d'une série [181][182].

- [181] Les problèmes que sont les incivilités, la petite délinquance, les atteintes non homicides aux personnes et les « violences urbaines » (émeutes dans les quartiers pauvres) apparaissent comme durables et croissants. Les statistiques policières sur la base desquelles sont estimés les taux d'incidence de ces problèmes ont été en France ou à l'étranger soumises à des critiques répétées. Elles présentent en effet de nombreuses limites (tous les faits ne sont pas déclarés par les victimes, suivant les types de délits, les propensions à la plainte ne sont pas uniformes, tous les faits déclarés ne sont pas enregistrés par la police notamment). Toutefois, deux éléments

militent dans le sens de la prise au sérieux les tendances lourdes obtenues à partir des données policières. Premièrement, tous les pays européens de l'Ouest ont connu des tendances comparables pour ce qui est des vols et des agressions depuis 1945. On ne peut donc réduire les tendances à un comportement d'enregistrement d'une police ou à une politique gouvernementale. Deuxièmement, le développement d'outils nouveaux servant à mesurer la pression de la délinquance sans passer par le filtre policier, les enquêtes de victimation, a permis de produire des données indépendantes (Insee, International Crime Victimization Survey, British Crime Survey). Il s'avère que les tendances statistiques décrites par les enquêtes sur les victimes convergent avec les chiffres de la police (aussi bien pour les enquêtes spécifiques à un pays, le Royaume-Uni, que pour les pays participant aux enquêtes européennes sur les victimes ou ICVS, notamment en France. Enfin, les rares études sur les violences anti-policières ou émeutes en France montrent le développement de ces explosions bien avant les émeutes les plus spectaculaires qui se sont pour la première fois étendues à la France entière en 2005. En résumé, au moment de la réforme, les éléments chiffrés dont on dispose indiquent une augmentation croissante de la pression de la délinquance contre les biens et les personnes sur près de quarante années (1945-1997), et une augmentation des phénomènes d'émeutes sur près de vingt années (1981-1997). [spo-art-241-body]

- [182] [...] Enfin, l'ajustement anticipé du mouvement au poids dépend progressivement de la taille et de la familiarité des objets. L'enfant parvient à prévoir le poids d'un objet jamais soulevé en opérant une prédiction sur la relation entre la taille et le poids. Ainsi, l'ajustement anticipé au poids dans une tâche de sériation apparaît nettement autour de 3,5 ans (Hauert, 1980 ; Hauert et al., 1981 ; Gordon, Forssberg, Johansson, Eliasson et Westling, 1992) et s'affine encore jusqu'à 9 ans (Gachoud, Mounoud et Hauert, 1983). En résumé, la capacité à extraire, stocker et utiliser par anticipation des propriétés des objets qui ne sont pas accessibles par la vision résulte d'un processus complexe dont le développement nécessite une dizaine d'années (Forssberg, 1999). [psy-art-579-body]

Par rapport aux marqueurs de clarification (*c'est-à-dire, à savoir, en d'autres termes, autrement dit*), le contexte gauche de *en résumé* est assez large et englobe une succession d'idées. Dans l'exemple [181], le locuteur développe dans la séquence **p** différents arguments qui relèvent des limites des *statistiques policières sur la délinquance et sur les émeutes*, ainsi que d'autres recherches et enquêtes menées indépendamment avec la police. Tous ces développements d'idées permettent au locuteur de conclure l'état de la délinquance et des émeutes à différentes périodes.

L'emploi de *en résumé* est identique dans l'exemple [182] où la séquence **p** (**p en résumé**) est composée d'un ensemble d'arguments afin de débattre sur la *capacité à extraire, stocker et utiliser par anticipation des propriétés des objets*. La reformulation établie par *en résumé* résulte de tout ce processus d'analyse et sert à souligner la conclusion saillante à la suite des arguments.

La séquence reformulatrice introduite par *en résumé* peut aussi correspondre à une généralisation par induction [183][184].

- [183] En France, nous avons vu des pratiques apparemment contradictoires : d'un côté la mitraille de questions et réponses rapides adressées à toute la classe, séquences dans lesquelles beaucoup d'enfants étaient impliqués, même brièvement, mais où il n'y avait pas de dialogue. De l'autre côté, de véritables dialogues soutenus du genre de ceux auxquels nous avons assisté encore plus souvent en Russie. Dans le Michigan, la façon de parler était détendue, conversationnelle et totalement non structurée. À l'évidence, cela donnait aux enfants confiance en eux-mêmes, ils ne montraient aucun signe d'intimidation comme le font quelquefois les enfants français. En revanche, en dehors de cela, aucun effort n'était requis des enfants, car le but était beaucoup plus social que cognitif. Bien sûr, il y a des exceptions, et l'enseignante du Michigan dont nous avons parlé, qui avait l'organisation la plus flexible que nous ayons trouvée, réussissait aussi à combiner harmonieusement un mode de parler ouvert et détendu qui s'avérait aussi structuré et intentionnel. En résumé, la parole pédagogique peut apparaître dans des contextes organisationnels divers : l'enseignement frontal, le travail collaboratif ou collectif en petits groupes, ou l'instruction individuelle face à face. Elle peut avoir des fonctions communicatives diverses aussi : interrogation, exposition, évaluation, dialogue ou même conversation. Et elle peut servir à instruire, à discuter, à évaluer, à partager l'information, à résoudre des problèmes, à « échafauder » connaissance personnelle et unique de l'enfant et connaissance des disciplines. L'apprentissage efficace a besoin de toutes ces sortes de paroles : en effet, il faut penser en termes de répertoire de l'enseignement, plutôt que de méthode. [sed-art-417-body]

Dans cet exemple, le contexte gauche de *en résumé* est tellement large que nous ne pouvons pas le citer ici. En effet, l'auteur y présente le rôle de la parole pédagogique au sein de l'école et met en comparaison différentes pratiques pédagogiques en France, en Russie et aux Etats-Unis. *En résumé* se trouve dans un nouveau paragraphe et en position détachée par rapport aux idées développées. Sa portée *q* est assez large et constituée de plusieurs propositions. Chaque proposition correspond successivement à une formulation plus générale des éléments que l'auteur a précédemment analysés.

Dans l'exemple [184], à l'instar de [183], la séquence reformulatrice introduite par *en résumé* se trouve dans un paragraphe séparé. La reformulation consiste en une généralisation des segments antérieurs qui s'opère par induction. L'auteur commence par une petite définition du *mood*. S'ensuivent différents exemples concrets sur la manière de le mesurer, pour arriver à la conclusion de sa fonction. La reformulation sert à relever le rôle essentiel que joue *le mood* dans cette étude. En raison de la taille des exemples et de l'analyse que le locuteur a effectués, nous ne pouvons que citer dans l'exemple ci-dessous une partie de l'étude en l'occurrence.

- [184] Une fois cette reconstitution effectuée, le logiciel crée une nouvelle série, le *mood* (ou l'indice longitudinal de préférence) en utilisant les valeurs de chacune des séries utilisées. Cette création est un processus itératif, en ce qu'il est renouvelé plusieurs fois afin de « coller » le mieux à la réalité des données : ce processus s'arrête une fois que l'on a atteint un maximum en termes ici de corrélations entre l'indice et chacune des séries qui le constituent. À noter que ce processus se déroule à la fois à rebours (ici de 2008 vers 1999) et chronologiquement (de 1999 à 2008) doublant ainsi les

méthodes d'estimation. *Enfin*⁴⁴, est estimée la corrélation entre l'indice longitudinal et chacune des séries intervenant dans sa constitution [...]. Face à ces évolutions qui, pour la plupart, se situent dans la marge d'erreur des sondages de cette taille d'échantillon, l'indice longitudinal enregistre une variation de + 1,5 point. Autrement dit, pour que l'indice change, les mouvements du public doivent être importants et s'opérer sur plusieurs séries. Cette inertie du *mood* est bénéfique à l'analyse, l'augmentation notable du soutien au droit de vote des étrangers pouvant fort bien être le produit d'un de ces aléas dont nous mentionnions précédemment l'existence. En résumé, le *mood* permet donc bien de prendre en compte les variations de l'opinion dans le temps sur une dimension latente, ici la tolérance aux étrangers, en mobilisant des séries incomplètes. Reste maintenant à profiter pleinement de la richesse du baromètre de la CNCDH. [spo-art-117-body]

Nous constatons donc que la séquence **p** suivie de *en résumé* est beaucoup plus volumineuse que celle de *en somme*. Elle peut se composer de différentes idées et nous relevons également d'autres marqueurs d'enchaînement thématique dans la séquence **p**. L'emploi de *en somme* à la place d'*en résumé* ne fonctionne pas dans les exemples [184] et [183]. Dans [183], le locuteur met en place une comparaison entre les deux systèmes avec des arguments parfois opposés, tandis que comme nous l'avons vu précédemment, *en somme* ne permet de relier que des éléments concordants. Dans [184], étant donné que le contexte gauche du MD est trop volumineux, il ne peut pas être suivi de *en somme*.

À la différence d'*en somme*, *en résumé* peut articuler des éléments concordants ainsi que des éléments contradictoires [185].

- [185] Par exemple, Amiot observe que la plupart des verbes actuels préfixés en pré- présentent un sens d'antériorité temporelle. Le verbe préchauffer correspond grossièrement à l'idée de # « chauffer au préalable, à l'avance ». Donc tout verbe présentant un sens d'antériorité temporelle sera considéré comme sémantiquement régulier. Par ailleurs, l'auteur relève les verbes prédominer ou prévaloir où le préfixe n'a pas une valeur d'antériorité temporelle, mais a un sens de supériorité. Cette propriété sémantique n'étant partagée que par ces deux verbes, elle est considérée comme une irrégularité au sein des verbes préfixés en pré-. En résumé, le degré de régularité d'une donnée se mesure à la quantité de mots dans l'ensemble des mots affixés partageant cette propriété. Les propriétés sémantiques d'un mot affixé seront considérées comme régulières parce qu'elles correspondent aux propriétés partagées par le plus grand nombre des mots comportant cet affixe. [lin-art-490-body]

Dans cet exemple, l'auteur précise avant tout le sens général des *verbes préfixés en – pré*, dont le sens d'antériorité temporelle sauf les deux verbes qui ont un sens de supériorité. *En résumé* y introduit une reformulation de valeur récapitulative qui repose sur la règle déduite dans le contexte précédent. L'emploi de *en somme* y est impossible en raison de l'apparition des verbes *prédominer* et *prévaloir* qui n'ont pas de sens d'antériorité temporelle.

⁴⁴Le marqueur est mis en italique par nous-même.

Nous essayons de représenter dans le Tableau 5.7 : Récapitulatif des propriétés syntaxiques et sémantiques de *en résumé* :

Paramètres	Propriétés
Position	Initiale. <i>En résumé</i> est toujours placé en tête de phrase, parfois il se trouve en position détachée pour introduire un paragraphe.
Portée	Globale. Les séquences p et q sont toujours de nature propositionnelle. Elles peuvent même se composer de plusieurs paragraphes. $\left(\begin{array}{c} \text{proposition} \\ \text{paragraphe} \end{array} \right)_p \text{ En résumé } \left(\begin{array}{c} \text{proposition} \\ \text{paragraphe} \end{array} \right)_q$
Valeurs sémantiques	✓ <i>Valeur épistémique</i> : <i>en résumé</i> sert à introduire une formulation plus condensée à la suite des idées évoquées dans la séquence p . $[\text{proposition}_1 + \dots + \text{proposition}_n / \text{paragraphe}]_p \text{ En résumé } [V.EPISTEMIQUE] [\text{proposition/paragraphe}]_q$

Tableau 5.7 : Récapitulatif des propriétés syntaxiques et sémantiques de *en résumé*

Pour conclure, *en résumé* sert à introduire une séquence à valeur épistémique dans l'objectif de tirer l'idée saillante pour le lecteur. Au plan sémantique, si *en somme* articule uniquement des éléments de même orientation argumentative, *en résumé* met en relation des éléments qui peuvent être contradictoires. Le premier a la propriété de souligner le cheminement de la pensée de l'auteur, tandis qu'*en résumé* établit une reformulation d'une manière plus objective, plus neutre qui résulte en général d'une série d'idées ou éléments. Au plan syntaxique, la taille de **p** qui précède *en résumé* est plus large que celle qui précède *en somme*. Elle peut être composée de plusieurs paragraphes, tandis que la séquence **p** de *en somme* est plus petite.

5.4.3. En un mot

a. Définition

Même si la fréquence de *en un mot* est relativement faible dans notre corpus (14 occurrences dont le ratio est de 0,05 par million de mots), il est encore un peu plus fréquent dans le corpus SHS que dans le corpus Emolex du français général (70 occurrences, avec un ratio de 0,02 par million de mots). De plus, dans le corpus Emolex, nous relevons plusieurs occurrences où *en un mot* n'est pas utilisé en tant que MD, par exemple : pour *le dire en un mot, expliquez-moi ce phénomène en un mot*, etc. Il nous paraît donc intéressant d'analyser son fonctionnement syntaxique et sémantique et de le comparer avec les autres marqueurs de même groupe. Comme les autres marqueurs, sa signification est assez transparente. Il sert en premier lieu à résumer ce qui a été dit et a pour synonyme *bref, pour résumer*.

b. Portée syntaxique

En un mot porte en général sur une séquence de nature propositionnelle (**P_G**) [187][188], mais il peut également prendre sous sa portée un syntagme (**P_L**) [186]. La taille de la séquence **p** varie selon que *en un mot* a une portée globale ou locale. Quand ce marqueur a une portée globale, la séquence **p** est aussi de nature propositionnelle [188]. Quand il a une portée locale, **p** correspond à un syntagme ou à une proposition [186].

- [186] Le projet de loi, muni de la signature royale et du contresceau ministériel, est alors déposé devant l'une ou l'autre des chambres. Le principe de la solidarité gouvernementale, fondé sur la partitocratie, garantit en général la durabilité du compromis. Celui-ci repose sur la loyauté, l'un des critères majeurs d'appréciation du personnel gouvernemental, comme le souligne A. Eraly : « Au-delà des rangées de fidèles et de sympathisants, s'étend le monde des partisans de la coalition avec lesquels il s'agit d'entrer dans toutes sortes d'échanges, de concertation et de négociations. Ce n'est plus ici la loyauté à l'homme et au parti, mais la loyauté aux accords, le respect de la parole et de la confidentialité, **en un mot la fiabilité**, qui deviennent la condition, d'autant plus valorisée qu'elle est fragile, du pouvoir politique. [...] Le respect de la confidentialité par exemple, finit par apparaître comme un événement rare et précieux, et comme la condition du succès des négociations. » [spo-art-142-body]

c. Position syntaxique

En un mot se trouve toujours en position initiale [187] et introduit la séquence reformulatrice **q**. Quand il porte sur une proposition, il est placé en tête de phrase. Il est en position postverbale quand il porte sur un syntagme.

- [187] Bref, la sociologie du droit comparée ne sera pas en mesure de rendre explicites les différences entre les pays, les auteurs, les mouvements et les écoles de pensée juridiques. En un mot, elle ne sera pas capable de comparer tant qu'elle n'aura pas accepté de prendre en compte certaines présuppositions liées au contexte social et culturel dans lequel ces différences se déploient. Voici deux de ces présuppositions[soc-art-228-introduction]

d. Valeurs sémantiques

Par son sémantisme, *en un mot* sert à introduire une formulation plus synthétique des segments précédents. La séquence reformulatrice a une *valeur résumante* [188].

- [188] Cette stratégie se développe particulièrement avec trois groupes clés : les élites politiques qui voient leur hégémonie et leurs intérêts menacés par de nouveaux groupes sociaux et prétendent revendiquer un surcroît de démocratie tout en en contournant les procédures habituelles ; les élites économiques qui perçoivent la constitutionnalisation des droits, notamment ceux concernant la propriété, la mobilité et les droits de l'emploi comme un moyen de peser sur l'action du gouvernement et de promouvoir un marché sans contrainte (a free market) et un agenda favorable aux affaires ; les élites judiciaires et les tribunaux supérieurs au niveau national enfin, lesquels cherchent à renforcer leur influence politique et leur réputation internationale. En un mot, selon Ran Hirsch, ce sont « des innovateurs stratégiques dans le domaine juridique (strategic legal innovators) [c'est-à-dire] les élites politiques en association avec les élites économiques et judiciaires qui ont des intérêts compatibles [qui] déterminent le timing, l'ampleur et le contenu des réformes constitutionnelles » (Hirschl, 2004, 12). [soc-art-35-body]

Dans cet exemple, le locuteur développe dans la séquence **p** le rôle essentiel de chacun de trois groupes *élites sociaux*. La portée **q** de *en un mot* correspond à une citation où est résumée l'importance de l'association de ces trois groupes. Les éléments de la séquence **p** concordent et vont dans le même sens que la séquence **q**. Le marqueur *en un mot* instaure un rapport plus ou moins objectif entre la première formulation et la séquence reformulatrice.

La fonction de *en un mot* est la même dans les exemples [189][190]. Cependant, dans ces exemples, le marqueur établit une démarche de généralisation entre la séquence **p** et **q**. *En un mot* a pour rôle d'orienter des éléments ou des points de vue plus spécifiques vers un élément plus général.

- [189] Compte tenu de sa complexité qui l'inscrit dans la ligne d'un fait social total, la baraka arabe pourrait être rapprochée de la *charis* grecque. À l'instar de celle-ci, la baraka représente une puissance divine, une force mystérieuse présente dans les mots, les choses et les êtres. Et comme la *charis*, le signifiant baraka s'applique à toute une série de notions associées, tels le don, la fertilité, la fécondité, la croissance, la prospérité, la santé, la beauté, la chance, le pouvoir, le charisme. En un mot, la notion de baraka symbolise la manifestation du divin dans le monde et l'existence, sous forme de dons accordés à certains individus (ou à certaines choses). [ant-art-347-body]

- [190] Le rapport aux ruines, en cela, a considérablement évolué en quelques décennies : « La tradition romantique nous a habitués à une vision des ruines où le rêve l'emporte

sur la réalité. Les paysages d'Hubert Robert, les récits des voyageurs-archéologues du xix^e siècle ont conforté cette image. Aujourd'hui, le goût de l'histoire, la vulgarisation des travaux historiques, la passion pour les "chefs-d'œuvre en péril" ont fait naître une autre approche de la ruine. Le visiteur veut comprendre ce qu'il voit, identifier la fonction des monuments dont ne subsistent plus que des fondations ; **en un mot, dépasser l'impression sentimentale pour atteindre la compréhension** » (Gauthier, Boiret, 1991, p. 31). Les ruines interrogent encore plus l'observateur lorsqu'il s'agit de ruines violentes en milieu urbain, et qui contrastent avec l'aménagement du reste de la ville. Le visiteur s'arrête devant la ruine et veut comprendre son histoire ; il peut être amené à réfléchir au risque et à la catastrophe. [geo-art-349-body]

Dans l'exemple [189], le locuteur explique dans la séquence **p** ce que représente *la baraka arabe*, ses propriétés essentielles en comparaison avec *la charis grecque*. Il le résume dans la reformulation par des notions plus générales. Dans [190], la reformulation établie par *en un mot* sert plutôt à nommer les idées précédentes. Dans ces exemples, le locuteur cherche à donner dans la séquence reformulatrice **q** les notions plus génériques et englobantes des idées évoquées dans la séquence **p**. Au plan syntaxique, *en un mot* met en relation deux séquences **p** et **q** de même nature syntaxique. Dans [189], *en un mot* a une portée globale, la séquence **p** est composée de plusieurs propositions. Dans [190], ce marqueur porte sur un syntagme verbal ainsi que dans la séquence **p** (**p** = *veut comprendre ce qu'il voit, identifier la fonction des monuments dont ne subsistent plus que des fondations*). Dans ces exemples, *en un mot* ne peut être remplacé ni par *en résumé* ni par *en somme*. En effet, sa portée est plus restreinte par rapport à celle de *en somme* et *en résumé*. Il aborde une seule idée dans la séquence **p**, tandis que deux autres marqueurs permettent de relier plusieurs idées.

Nous pouvons résumer les propriétés principales de *en un mot* dans le Tableau 5.8 : Récapitulatif des propriétés syntaxiques et sémantiques de *en un mot*.

Paramètres	Propriétés
Position	Initiale. Au niveau phrastique, il est placé en tête de phrase quand <i>en un mot</i> porte sur une proposition et en position postverbale, quand le marqueur porte sur un syntagme.

Portée	<p>Globale et locale. La portée q de <i>en un mot</i> est régie par la nature syntaxique de la séquence p.</p> <div style="text-align: center;"> $\left(\text{phrase} \right)_p \text{ En un mot } \left(\text{phrase} \right)_q$ $\left(\left(\text{syntagme nominal} \right) \left(\text{syntagme verbal} \right) \right)_p \text{ en un mot } \left(\left(\text{syntagme nominal} \right) \left(\text{syntagme verbal} \right) \right)_q \text{ phrase}$ </div>
Valeurs sémantiques	<p>✓ <i>Valeur résumante</i> : la reformulation correspond à une formulation plus synthétique du contenu sémantique de la séquence p.</p> <p>➤ Quand il a une <u>portée globale</u> :</p> <p>[proposition]_p En un mot V.RESUMANTE [proposition]_q</p> <p>➤ Quand il a une <u>portée locale</u> :</p> <p>[SN/SV]_p en un mot V.RESUMANTE [SN/SV]_q</p>

Tableau 5.8 : Récapitulatif des propriétés syntaxiques et sémantiques de *en un mot*

En conclusion, l'emploi de *en un mot* se rapproche de *en somme* dans le sens qu'elle peut articuler uniquement des éléments concordants. Pourtant, ce marqueur envisage moins un mouvement de synthèse qu'*en somme*, mais plutôt une démarche de généralisation des idées ou points de vue antérieurement développés. Parmi ces trois marqueurs récapitulatifs, *en résumé* permet d'établir un rapport plus objectif entre la séquence **p** et **q**. Sur le plan syntaxique, *en un mot* est utilisé dans des contextes plus restreints qu'*en résumé* ou *en somme*. En fait, le contexte gauche de ces deux marqueurs est constitué en général d'une série d'idées et de points de vue et peut s'étendre à plusieurs paragraphes, tandis qu'on relève une seule idée principale dans la séquence **p** de *en un mot*.

Nous mettons en évidence quelques propriétés saillantes qui caractérisent l'emploi de chaque MD de cette catégorie dans le Tableau 5.9 : Comparaison des caractéristiques de *en somme*, *en résumé* et *en un mot*.

Marqueurs discursifs	Paramètres
En somme	<ul style="list-style-type: none"> - Il est utilisé après avoir passé en revue tous les éléments et indique le cheminement de la pensée du locuteur. <i>En somme</i> marque un changement d'angle de vue pour faire progresser le raisonnement. - <i>En somme</i> met en relation uniquement les éléments ou idées concordants. - Il porte essentiellement sur une proposition, mais il peut éventuellement porter sur un syntagme.
En résumé	<ul style="list-style-type: none"> - Il est utilisé pour relier des idées concordantes ou contradictoires. - <i>En résumé</i> introduit des reformulations plus neutres, plus objectives. - <i>En résumé</i> est remplacé par <i>en somme</i> uniquement quand la séquence p introduit des unités qui sont concordantes entre elles. - La séquence p est en général très large et peut se composer d'un ou même plusieurs paragraphes. Elle est plus grande que celle de <i>en somme</i>. - La portée q peut correspondre à un paragraphe. <i>En résumé</i> peut être placée en position détachée pour introduire un nouveau paragraphe, ce qui n'est pas le cas de <i>en somme</i>.
En un mot	<ul style="list-style-type: none"> - Il ne permet de reformuler qu'une seule idée dans la séquence p. - La séquence p est moins grande que celle de <i>en somme</i> et <i>en résumé</i>. - La reformulation établie par <i>en un mot</i> correspond plutôt à une généralisation, tandis qu'<i>en somme</i> et <i>en résumé</i> envisagent d'effectuer une synthèse.

Tableau 5.9 : Comparaison des caractéristiques de *en somme*, *en résumé* et *en un mot*

5.5. Reconsidérer

Nous rejoignons Lenepveu (2008) sur la définition de l'opération de la reconsidération qui suppose que « le locuteur a passé en revue tous les éléments censés justifier une première formulation (implicite ou explicite) et qu'à partir de cette reconsidération, le locuteur est à

même d'introduire un nouveau point de vue » (*ibid.* : 113). Nous en distinguons deux groupes : 1/ *au total*, *somme toute* portent sur l'analyse de la totalité des éléments, 2/ *en définitive*, *en fin de compte*, *au fond* insistent sur le point de vue donné par l'auteur dans la séquence reformulatrice comme ultime.

5.5.1. Au total

a. Définition

La signification de *au total* est liée à son composant *total*. Quand *au total* est un adverbial quantificateur, il permet d'introduire une addition de nombres. Dans ce cas, il est utilisé avec les termes relatifs aux quantités ou à l'établissement des rapports inférieurs ou supérieurs [191][192][193].

- [191] Le Département affirme avoir audité 7 612 cybercafés **au total** durant cette période. Il en a fermé 1 068 opérant sans licence, condamné à des amendes 635 cybercafés qui violaient les règles opérationnelles et mis sous scellés 630 cybercafés en marge de la réglementation. [geo-art-34-notes]
- [192] Le pays d'origine est le Maroc pour 12 d'entre eux, la Turquie pour 2 autres, et les autres étaient originaires de la Tunisie, du Congo et de la Guinée. Sur les 17 élus flamands du Parlement régional bruxellois, 1 est d'origine marocaine. **Au total**, 20,2 % des 89 élus sont d'origine non européenne. [spo-art-153-body]
- [193] Sur la base de notre compétence linguistique, nous avons dressé manuellement une liste exhaustive des énoncés définitoires contenus dans chacun des textes du corpus. Cette liste qui rassemble **au total** 1 574 énoncés définitoires est indispensable pour faire les mesures statistiques permettant d'évaluer l'efficacité des structures linguistiques pour le repérage automatique des segments textuels où un énoncé définitoire est exprimé. [lin-art-169-body]

Ces trois exemples illustrent trois cas de figure de l'emploi de *au total* pour exprimer une quantification. Si *au total* dans [191] est un simple adverbial quantificateur, les séquences introduites par ce marqueur dans [192] et [193] sont en effet une autre formulation des séquences antérieures. Cet emploi de *au total* en tant qu'adverbial quantificateur nous intéresse moins. Nous excluons de notre étude tous les cas où *au total* introduit une quantification [191] ou une addition équivalente grâce aux différents modes de calcul [192][193] et ne conservons que le deuxième sens de *au total*, quand celui-ci indique un processus d'analyse de l'auteur. Nous retenons 49 occurrences de *au total* comme un marqueur de reformulation (sur le total de 79 occurrences dans notre corpus). D'après le TLFi, *au total* a pour synonyme *en fin de compte*, *tout compte fait*, *en somme*.

b. Portée syntaxique

En raison de son sémantisme, *au total* ne prend sous sa portée qu'une proposition (P_G) [194]. Son contexte gauche est également de nature propositionnelle et se compose de plusieurs propositions.

- [194] Cependant, dans les autres cas, l'absence de culture nouvelle, voire de nouveaux acteurs, n'induit pas nécessairement une faible activité du marché foncier et les producteurs de riz, de canne à sucre et de café tentent aussi d'agrandir leurs exploitations. La clef de la diversité constatée est l'importance variable des terres non cultivées de propriété collective et leur accessibilité pour les investisseurs potentiels. Celle-ci dépend de caractéristiques physiques ou spatiales (proximité des routes, climat, etc.), mais également des statuts fonciers préexistants, résultat d'histoires agraires complexes, qui sont plus ou moins aisément modifiables par la législation libérale des années 1990. Au total, ces territoires bien individualisés réagissent différemment au nouveau contexte économique et juridique. [geo-art-1-summary]

c. Position syntaxique

Au total se trouve en position initiale et introduit la séquence reformulatrice. Il peut être précédé d'un autre marqueur, par exemple dans [195] où il est précédé par *bref*. En tête de phrase, *au total* est suivi d'une marque graphique dont la virgule. Il peut être placé en position postverbale [196].

- [195] Les populations locales s'organisent, se regroupent, spontanément, à l'instigation de groupes internationaux, etc. Pour l'instant, notre analyse n'a pas encore examiné cet aspect des transformations. Mentionnons seulement un exemple, celui d'un réseau de Karen en Thaïlande, fort bien structuré, le *Karen Network for Culture and Environment* [20].

Bref, *au total*, loin d'y perdre au change, du moins jusqu'à maintenant, l'État demeure bien, comme le précisais Wai-Chung Yung (1998, p. 294), l'acteur le plus important de la vie quotidienne, avec bien sûr des variantes importantes. L'État a besoin d'une population, d'un territoire et d'une relation de pouvoir pour les relier à une autorité (Raffestin, 1980 ; Déry, 2004). [geo-art-134-conclusion]

- [196] Mais cette vie de quartier « traditionnelle », associant relations et usages de proximité, occupe *au total* une place assez limitée dans leur vie sociale, qui apparaît fortement structurée autour de deux univers nettement dissociés dans le temps et dans l'espace : le travail (la semaine, à Paris) et la famille (le week-end, à Versailles). [soc-art-157-body]

d. Valeurs sémantiques

D1. Valeur épistémique

Au total peut introduire une séquence à valeur épistémique qui consiste en un jugement de l'auteur à la suite d'une série d'éléments. *Au total* intervient au terme d'une opération ou d'une description d'éléments assez implicites [197][198].

[197] Ces résultats demandent à être confirmés et approfondis. *En premier lieu*⁴⁵, ils sont obtenus sous une hypothèse paramétrique particulière (la normalité des résidus du modèle) qui mériterait certainement d'être assouplie. *Ensuite*, l'équation de sélection dans le dispositif de formation continue ne tient pas compte du salaire préalablement perçu par l'employé, et cette omission peut avoir des conséquences sur les résultats d'estimation. *Enfin*, nous n'avons pas tenu compte de la durée des formations reçues, ni de leur contenu. De telles informations existent dans l'enquête FQP, mais la taille de l'échantillon ne semble pas suffisante pour estimer précisément les différences d'effets qui pourraient leur être associées. **Au total, si la formation continue a peu d'effet sur les carrières individuelles, cela ne signifie pas qu'elle n'en a pas sur la situation des entreprises formatrices.** [eco-art-22-conclusion]

Ce paragraphe est selon nous assez bien structuré. *Au total* ne souligne pas la somme des arguments analysés. Il met en relation les arguments avec un point de vue. Le lecteur assiste au processus de construction de la pensée de l'auteur. Cet extrait peut correspondre à la partie de discussion dans un écrit scientifique où le locuteur évalue dans la séquence **p** les résultats qu'il a précédemment obtenus. Il donne dans la séquence **q** son jugement sur la conclusion finale qui traite de l'influence de la formation continue sur *la situation des entreprises formatrices*. Ce qui distingue *au total* de *en somme* est le fait que le premier permet même de rassembler les arguments contradictoires du contexte gauche, tandis que le deuxième sert à reformuler uniquement les idées concordantes.

Les exemples [197] et [198] ont le même schéma textuel. Ils sont construits sur le principe d'un raisonnement inductif. Le locuteur part de son propre constat du contexte actuel [198] ou de différents résultats [197] pour arriver à donner son propre point de vue dans la séquence **q** introduite par *au total*. Nous rejoignons Roulet (1987) et Rossari (1990) pour affirmer que *au total* introduit un changement de perspective. Comme dans [198], l'auteur détaille dans la séquence **p** les changements du paysage social et démographique. La séquence reformulatrice **q** correspond à un jugement du locuteur de l'influence de ces changements sur la cohésion sociale.

[198] Les influences de cette mixité entre villageois résidents, migrants venus du delta et citadins, créent un paysage social et démographique à facettes multiples. Les jeunes ménages de la classe moyenne, notamment, accèdent plus largement à la location ou à

⁴⁵ Les marqueurs d'enchaînement thématique sont mis en italique par nous-même.

l'achat de logement (dans des ensembles d'habitat collectif ou, pour les plus aisés, dans des villas) dans ces mêmes zones ou dans les nouvelles zones urbaines. Parmi les mutations sociales majeures générées par la croissance spatiale de la ville, la reconversion des paysans est l'une des plus radicale. En effet, l'expropriation des terres agricoles au profit de l'expansion des nouveaux quartiers, autrefois ponctuelle et très localisée, se généralise sur les franges urbaines, et pose la question de la reconversion socioprofessionnelle des paysans. **Au total, si la transition urbaine s'accompagne aujourd'hui de recompositions territoriales, tant à l'intérieur de la ville que sur les espaces périurbains, elle interroge donc aussi la cohésion sociale.** [geo-art-6-body]

Dans l'exemple [199], *au total* n'envisage pas d'apporter une conclusion sur l'état de l'aide multilatérale, mais uniquement l'opinion de l'auteur vis-à-vis de la participation des *particuliers qui transitent par les ONG*. Dans cet exemple, *au total* articule deux idées contradictoires, *un grand nombre de dons d'un côté et la somme de participation relativement faible d'un autre côté*.

[199] L'aide multilatérale provient soit des contributions venant des budgets des pays riches, soit des particuliers que par exemple achètent les cartes de vœux de l'UNICEF. Il y a enfin les dons des particuliers qui transitent par les ONG, qui existent en très grand nombre, mais qui **au total ne contribuent qu'à hauteur de quelques pourcents aux flux d'aide.** [sed-art-464-body]

Dans [200][201], l'emploi de *au total* joue le même rôle que dans les exemples précédents où le marqueur introduit le point de vue de l'auteur.

[200] Néanmoins, l'approche en coupe transversale ne prend pas en compte l'historique des revenus de chaque génération. Or, au-delà du niveau de vie atteint à un instant donné, la croissance régulière de celui-ci est également un facteur de bien-être pour les individus. Ces derniers se sentent *par ailleurs* d'autant moins « lésés » par rapport à leurs contemporains que leur progression est parallèle à celles des autres générations aux mêmes âges. D'où l'intérêt d'une approche complémentaire en coupe longitudinale. **Au total, il apparaît donc nécessaire de prendre en compte, dans l'appréciation de l'équité du système de retraite entre les générations, non seulement les différences instantanées de niveau de vie, mais également les différences historiques aux âges comparables.** [eco-art-608-body]

[201] Le statut administratif embryonnaire a d'abord évolué en fonction des liens de vassalité qu'entretenaient certains des groupes avec l'extérieur (Viets ou Cham). Ce statut a changé avec l'arrivée plus formelle des Français dans cette région à la fin du XIX^e siècle, les explorateurs et administrateurs suivant les missionnaires. Les premières unités administratives ont été délimitées au tout début du XX^e siècle. Modifiées plusieurs fois jusqu'au départ des colonisateurs, ces limites et les unités qu'elles délimitaient sont demeurées, **au total, peu nombreuses, témoignant davantage de la carte des densités de population que de celle des catégories ethniques établies.** [geo-art-134-body]

Dans l'exemple [200], l'auteur explique de prime abord l'intérêt de deux approches dont *l'approche en coupe transversale* et *l'approche en coupe longitudinale*. Ces différents arguments vont dans le sens de la conclusion visée et sont conformes à la suite logique du raisonnement quand l'auteur plaide pour l'association de ces deux méthodes. Dans [201], la

reformulation qu'instaure *au total* sert à introduire le jugement de valeur de l'auteur sur *les premières unités administratives au début du XX^e siècle*.

Dans ces exemples, *au total* ne peut être remplacé par aucun des marqueurs que nous avons abordés. Il permet d'introduire le point de vue et le jugement du locuteur, ce qui ne peut pas être véhiculé par des marqueurs d'autres catégories.

D2. Valeur résumante

La séquence reformulatrice établie par *au total* peut avoir une valeur résumante. Dans les exemples [202][203], la reformulation s'effectue par une démarche de généralisation et envisage plutôt de récapituler différents faits ou points de vue développés.

[202] [...] Le taux d'activité remonte ensuite puis recommence à baisser avec les cessations d'activité après 50 ans. Cette décroissance du taux d'activité liée à la maternité n'apparaît plus pour les générations postérieures à 1943, que cela soit dû aux interruptions moins fréquentes ou aux débuts plus tardifs. **Au total, les femmes des générations les plus récentes accumulent moins d'années de travail que leurs aînées avant 30 ans, mais tendent à les rattraper ensuite.** [eco-art-141-body]

[203] Pour le secteur mécanique, il faut noter la localisation croissante au sein de l'Allemagne et une diffusion presque inexistante vers les espaces périphériques, à l'exception de la périphérie nordique hautement spécialisée dans certains segments industriels très technologiques. En Europe centrale et orientale, il faut noter le déclin du secteur mécanique dans les zones les plus orientales comme la Bulgarie, la Roumanie ou les États baltes, ces derniers ayant par le passé été fortement spécialisés dans ces segments industriels qualifiés dans le cadre de la division du travail au sein de l'ancien système socialiste. Tabl. 2 - Indices de spécialisation des types économiques régionaux dans la chimie et le secteur mécanique entre 1980 et 2004. **Au total, l'évolution de la répartition de l'activité de quelques secteurs industriels au sein de l'espace européen confirme les logiques mises en avant à l'échelle mondiale.** [geo-art-80-body]

Dans ces exemples, les points de vue dans le contexte gauche et dans le contexte droit sont concordants. Dans l'exemple [202], la reformulation établie par *au total* correspond à une idée synthétique de ce qui a été développé dans la séquence **p**. Dans l'exemple [203], la reformulation établie par *au total* consiste en une synthèse de toutes les idées que le locuteur a développées dans le contexte gauche. Pourtant, l'emploi de *en somme* y est impossible, car la formulation synthétique est plus ou moins subjective et marquée par les traces de l'auteur *on*, *relativement*, etc., le verbe conjugué au conditionnel, etc.

Concernant le contexte gauche de *au total*, comme ce marqueur est utilisé à la suite des arguments et idées, la séquence **p** est toujours de nature propositionnelle et se compose de plusieurs propositions. Nous remarquons l'utilisation d'autres types de MD dans la séquence **p**, par exemple les marqueurs d'enchaînement thématique (*en premier lieu, ensuite, enfin*) dans [197], les marqueurs d'addition (*par ailleurs*, etc.) dans [200]. Ce point nous semble très

important pour l'enseignement/apprentissage des MD. Le repérage des MD pourrait aider les apprenants à relever des différentes structures discursives ainsi que les thématiques essentielles. Cet aspect sera abordé lors de la discussion sur les stratégies de lecture (Cf. Chapitre 6).

La portée **q** de *au total* peut être plus ou moins longue. Ce marqueur peut se trouver en position détachée pour introduire un paragraphe [204].

- [204] L'analyse de ces avis justifie cette étude en dernier lieu. Il importe d'apprécier en effet l'opinion des Français en distinguant leurs préoccupations financières. On voit alors apparaître, à côté des experts administratifs ou politiques, des groupes qui se prononcent sur la rente soit pour la défendre, soit pour la condamner. Que les attentes des capitalistes français l'emportent sur le ressentiment des patriotes ne témoigne-t-il pas à quel point la nation créditrice liait bel et bien son destin dans ces années décisives à celui de « l'État de finances » ?

Au total, nous espérons livrer une étude qui défende une approche sociopolitique de l'histoire des finances. L'analyse de toute gestion publique doit être indissolublement liée à celle des capitaux qui la porte certes, mais doit aussi donner en retour la mesure des effets de cette gestion sur son environnement. [his-art-166-body]

Nous récapitulons dans le Tableau 5.10 : Récapitulatif des propriétés syntaxiques et sémantiques de *au total* les propriétés principales de *au total*.

Paramètres	Propriétés
Position	Initiale. Il est toujours placé en tête de phrase au niveau phrastique, il peut se trouver en tête de paragraphe.
Portée	Globale. Il porte toujours sur une séquence de nature propositionnelle, soit une proposition soit un paragraphe. La séquence p correspond également à une proposition ou un paragraphe. <div style="text-align: center; margin-top: 20px;"> $\left(\begin{array}{c} \text{proposition} \\ \text{paragraphe} \end{array} \right)_{\mathbf{p}} \quad \mathbf{Au\ total} \quad \left(\begin{array}{c} \text{proposition} \\ \text{paragraphe} \end{array} \right)_{\mathbf{q}}$ </div>
Valeurs sémantiques	<p>✓ <i>Valeur épistémique</i> : <i>Au total</i> met en relation une suite d'idées et d'arguments avec le jugement d'auteur. Il indique le cheminement du raisonnement.</p> <p>[proposition₁+...+proposition_n/paragraphe]_p Au total V.EPISTEMIQUE </p>

	[proposition/paragraphe] _q
✓	Valeur résumante : la reformulation établie par <i>au total</i> correspond à une formulation plus ou moins subjective.
	[proposition ₁ +...+proposition _n /paragraphe] _p Au total V.RESUMANTE [proposition/paragraphe] _q

Tableau 5.10 : Récapitulatif des propriétés syntaxiques et sémantiques de *au total*

En bref, quand *au total* fonctionne comme un marqueur de reconsidération, il peut introduire des séquences de valeur épistémique et résumante. À la différence de *en somme*, *au total* peut articuler des éléments contradictoires et la reformulation qu'il instaure est empreinte de la subjectivité de l'auteur. L'auteur peut y apporter son jugement et son point de vue.

5.5.2. Somme toute

a. Définition

Roulet (1987) décrit *somme toute* comme connecteur de reformulation qui prend en compte la « totalité des éléments envisagés précédemment, voire envisageable » (Roulet, 1987 : 128) et il peut s'agir d'éléments positifs ou négatifs. La signification du marqueur vient du sémantisme de ses composants et signifie *à tout prendre* (d'après le TLFi). Il est également à noter que *somme toute* est peu fréquent dans le corpus Emolex français général (75 occurrences) dont le ratio est de 0,02 par million de mots contre 13 occurrences dans notre corpus SHS (ratio de 0,05). Certes, la différence de fréquence n'est pas importante, mais nous relevons quelques différences d'ordre syntaxique. En effet, dans le corpus Emolex, uniquement 5 occurrences portent sur un syntagme adjectival, 6 occurrences sur un syntagme nominal. Ils prennent tous sous leur portée une proposition, ce qui est l'opposé de ce que l'on observe dans notre corpus SHS.

b. Portée syntaxique

Somme toute porte essentiellement sur un syntagme adjectival (**P_L**) [205]. Deux occurrences sur le total de 13 occurrences portent sur une proposition (**P_G**) [206][210]. Quand

ce marqueur prend sous sa portée un syntagme adjectival, la séquence **p** correspond soit à un syntagme nominal [207] soit à une proposition faite avec les verbes d'état comme *demeurer*, *paraître*, etc. [205]. Le syntagme adjectival a une fonction qualificative.

- [205] Si une telle prodigalité demeure **somme toute** exceptionnelle et limitée essentiellement au XIII^e siècle, ce qui frappe plus encore est le souci des hommes d'approvisionner les indigents en victuailles, particulièrement en rations frumentaires vitales (30 /43 disposants), comme s'ils se faisaient l'écho plus ou moins lointain des litanies maintes fois émises au cours des délibérations municipales de l'époque, notamment à partir de la seconde décennie du XIV siècle. [his-art-63-body]
- [206] Avec cet ensemble d'efforts plus ou moins articulés, c'est **somme toute** un réseau d'initiatives qui tend à s'organiser et qui s'agence dans une forme de travail sur la complicité du consommateur citoyen, appelé à devenir « consom'acteur ». [spo-art-642-body]

c. Position syntaxique

Somme toute est en général placé en position initiale et introduit la séquence **q**. Concernant sa position au niveau phrastique, il se trouve en général en position postverbale [207] et nous relevons très peu d'occurrences où *somme toute* est placé en tête de phrase. [210].

- [207] Comme l'on pouvait s'y attendre dès lors que l'on considère le public cible d'une bibliothèque universitaire, la prise en main proprement dite des tablettes, si rustique soit leur ergonomie, ne pose pas de réel problème à de jeunes personnes habituées à utiliser quotidiennement des outils aux fonctionnalités autrement plus avancées que celles, **somme toute** élémentaires (pensez donc, lire !), d'un livre. [infcom-art-232-body]

d. Valeurs sémantiques

La séquence reformulatrice introduite par *somme toute* correspond au point de vue du locuteur sur l'élément précédent. Il s'agit d'un jugement à valeur épistémique [208][209][210].

- [208] Voisiner, c'est « tenir sa place » et laisser à l'autre la possibilité de tenir la sienne en toute souveraineté. De la même façon, la demande d'un service à un voisin sera jugée illégitime et abusive si la personne qui demande n'a pas la possibilité de restituer le service alors qu'elle a la ressource à portée de main pour se débrouiller par elle-même. C'est le cas d'une vieille dame âgée, dont la marche est difficile : elle vient demander à sa voisine d'aller pour elle à la pharmacie du quartier, alors que sa fille est auprès d'elle ; la voisine interrogée, plus « jeune », mais elle aussi à la retraite, fait le récit à l'enquêteur d'une telle demande, **somme toute** assez banale, mais qui la scandalise parce qu'« il y a de l'abus ». [soc-art-59-body]
- [209] Que penser, dès lors, de nos questions de départ, qui mettaient au centre du questionnement les enjeux de frontières ? Il semble moins évident maintenant que les

formes festives piémontaises ne font que déborder d'Italie depuis 1947 à Tende. Le chant est plutôt un phénomène complexe qui a finalement autant à voir avec l'histoire des frontières politiques qu'avec l'évolution récente des formes populaires de la musique vocale et l'histoire migratoire **somme toute** récente. [ant-art-222-body]

- [210] L'implicite ethnique réside dans le fait de nier la possibilité d'unions mixtes puisqu'aucune catégorie ne leur correspond. Les naissances d'unions mixtes sont actuellement plus nombreuses en France que les naissances issues de deux parents étrangers. On remarquera aussi que l'attribution a été effectuée par hétéro-évaluation puisque les enquêtés n'avaient pas le moindre mot à dire. Les responsables de l'enquête ont justifié leur choix en affirmant que si la question de l'appartenance avait été directement posée, les enquêtés se seraient très souvent identifiés comme Français. **Somme toute**, cela aurait répondu à l'objectif d'une enquête sur l'intégration (c'est dans l'intitulé de MGIS) et correspondu à la célèbre définition électorale de l'appartenance selon Ernest Renan. [ant-art-312-body]

Dans l'exemple [208], *somme toute* prend sous sa portée un syntagme adjectival qui indique l'opinion de l'auteur vis-à-vis de la demande de la voisine de la vieille dame âgée. L'apparition de l'adjectif *telle* (dans *telle demande*) renvoie à son contexte antérieur. À l'instar des marqueurs récapitulatifs, l'usage de *somme toute* demande à l'auteur d'avoir passé en revue tous les éléments précédents. Pourtant, si ces marqueurs établissent un rapport de généralisation en allant du plus spécifique au plus général, *somme toute* instaure un rapport peu commun entre la première formulation et la séquence reformulatrice - car il est caractérisé par la subjectivité de l'auteur. Nous rejoignons Roulet (1987) en confirmant qu'il opère un changement de perspective énonciative entre la séquence **p** et **q**. L'emploi de *somme toute* a la même valeur dans [210] quand ce marqueur porte sur une proposition. La séquence reformulatrice correspond au point de vue de l'auteur suite à la description sur *l'implicite ethnique et les naissances d'unions mixtes*.

Dans les exemples [211][212], la séquence **q** (dans *somme toute q*) sert à caractériser la série d'éléments dans la séquence **p**.

- [211] Lors de son arrivée, l'enfant adopté est encore situé entre deux familles, entre deux parentés ; il n'appartient plus à l'une, mais n'est pas encore intégré à l'autre. Il est, en quelque sorte, en marge et subit donc, comme toute personne associée à cet état frontalier et **somme toute** dangereux, les regards curieux, mais aussi méfiants et soupçonneux de tous ceux qui entourent la nouvelle famille. [ant-art-271-body]
- [212] Le problème est que cette description n'explique pas pourquoi une manière très indirecte de demander de passer le sel, comme (3), ne respecte pas les principes de politesse, et qu'une formule elliptique comme (4) est à la fois directe, explicite et très impolie: (3) La soupe manque de sel. (4) Le sel ! Si les explications en termes de contribution coopérative et de politesse ne sont pas satisfaisantes, comment expliquer une donnée fournie par (1), **somme toute** élémentaire et routinière ? [lin-art-509-body]

Dans ces deux exemples, *somme toute* porte sur un syntagme adjectival. En l'analysant au niveau phrastique, la suite reliée par le marqueur (**p somme toute q**) fonctionne comme un

syntagme adjectival dans lequel la séquence **q** joue le rôle d'un adjectif qualificatif. De plus, nous remarquons que la séquence **p** introduite par *somme toute* peut comporter des éléments grammaticaux comme des déterminants définis, des adjectifs (*tel* dans *telle demande* dans [208]) ou des pronoms démonstratifs (*cet* dans *cet état* [211]), ce qui explique que *somme toute* est utilisé uniquement à la suite d'un processus d'analyse préalable ou quand le contenu sémantique dans la séquence **p** est connu par le lecteur. En employant *somme toute*, l'auteur envisage moins de compiler une série d'éléments ou de points de vue, mais de les caractériser et de donner son propre point de vue. Dans [211], l'auteur souligne dans la séquence reformulatrice introduite par *somme toute* l'état inquiétant de l'enfant *situé entre deux familles*. Le marqueur a la même valeur dans [212] où l'auteur cherche à qualifier la valeur de la donnée dans un exemple précis.

Dans ces exemples, *somme toute* ne peut être remplacé ni par *au total* ni par des marqueurs d'autres catégories. Comme ce marqueur porte essentiellement sur un syntagme adjectival, l'utilisation de *au total* ne peut pas fonctionner. Les marqueurs récapitulatifs ne permettent pas d'introduire le jugement du locuteur.

En outre, nous constatons qu'en raison de la nature syntaxique de la séquence **q**, *somme toute* peut être précédé des verbes d'état comme *demeurer*, *paraître*, etc. [205][213].

[213] Sans critiquer le fait que Béatrice Vacher tienne à établir la différence entre ces dernières, il me semble intéressant de travailler le lien du couple info com avec l'organisation qui me paraît **somme toute** moins travaillé dans cet article, alors même que nombre de ses écrits prouvent le contraire. [infcom-art-311-body]

[214] Les processus d'unification des espaces publics électoraux demeurent **somme toute** assez faibles, aussi bien sur le plan juridique que politique. [spo-art-438-body]

Les propriétés essentielles de *somme toute* figurent dans le Tableau 5.11 : Récapitulatif des propriétés syntaxiques et sémantiques de *somme toute*.

Paramètres	Propriétés
Position	Initiale pour introduire la séquence q . Au niveau phrastique, <i>somme toute</i> se trouve souvent en position postverbale.

Portée	<p>✓ Locale (sur un syntagme adjectival). La séquence p correspond à un syntagme nominal ou à une proposition faite avec des verbes d'état <i>demeurer, paraître</i>, etc.</p> $\left(\begin{array}{c} \text{syntagme nominal} \\ S + V \text{ d'état (demeurer, paraître,} \end{array} \right)_p \text{ somme toute } \left(\begin{array}{c} \text{syntagme} \\ \text{adjectival} \end{array} \right)_q \text{ }_{\text{phrase}}$
Valeurs sémantiques	<p>✓ <i>Valeur épistémique</i> : la reformulation établie par <i>somme toute</i> sert à caractériser la séquence p et à donner le point de vue du locuteur.</p> <p>➤ Quand il a une <u>portée locale</u> :</p> <p>[syntagme nominal]_p somme toute V.EPISTEMIQUE [syntagme adjectival]_q</p> <p>➤ Quand il a une <u>portée globale</u> :</p> <p>[proposition]_p somme toute V.EPISTEMIQUE [proposition]_q</p>

Tableau 5.11 : Récapitulatif des propriétés syntaxiques et sémantiques de *somme toute*

En bref, *somme toute* peut se rapprocher de *en somme* sémantiquement. Cependant, *en somme* instaure, d'une part, un rapport de généralisation entre la première formulation et la séquence reformulatrice. Il porte en général sur une proposition et la séquence **p** (dans **p** *en somme*) est de nature propositionnelle. D'autre part, la séquence reformulatrice introduite par *somme toute* est empreinte de la subjectivité de l'auteur et porte en général sur un syntagme adjectival. L'emploi de ce marqueur résulte d'un processus d'analyse antérieur de l'auteur, ce qui s'explique par l'ajout d'éléments qui renvoient au contexte précédent (comme *tel*, etc.). En comparaison avec *au total* marqueur de reconsidération, *somme toute* et *au total* introduisent tous le point de vue de l'auteur. Pourtant, *au total* peut articuler même des éléments contradictoires, ce qui n'est pas le cas de *somme toute* qui permet d'articuler seulement des éléments concordants.

5.5.3. En définitive

a. Définition

En définitive fait partie des connecteurs reformulatifs non paraphrastiques dans le classement de Roulet (1987) (Cf. Chapitre 4, 4.3.1) ; il est considéré comme un des marqueurs de l'opération de réexamen par Rossari (1994). Selon Roulet (*op.cit.*), *en définitive* sert à introduire une reformulation considérée comme « le dernier d'une succession de points de vue, donc comme le point de vue définitif de l'énonciateur » (Roulet, *op.cit.* : 129). *En définitive* est légèrement plus fréquent dans notre corpus SHS (54 occurrences dont le ratio est de 0,19) que dans le corpus Emolex français général (110 occurrences avec un ratio est de 0,03 par million de mots). D'après le TLFi, il prend pour synonymes *tout bien considéré*, *en conclusion*.

b. Portée syntaxique

En définitive prend en général sous sa portée une proposition (**P_G**) [215] ou même un paragraphe [216]. La séquence **p** est aussi de nature propositionnelle et peut se composer de plusieurs propositions.

- [215] [...] De plus, leur spécialisation entraîne une désaffiliation du monde de la musique et de la société conventionnelle. Les musiciens de Mohamed Ali, en effet, sont spécialisés dans les musiques populaires urbaines, très demandées dans les mariages *baladi*, mais réprouvées par les tenants de la culture officielle et les promoteurs d'une esthétique musicale plus « civilisée ». Si les musiques labellisées « folklore » jouissent d'une reconnaissance sociale et institutionnelle en tant que témoignages de l'authenticité culturelle égyptienne, les musiques urbaines populaires, telles qu'elles se développent par le biais du marché de la cassette audio depuis une trentaine d'années, font, elles, l'objet d'une virulente stigmatisation dans les médias égyptiens comme dans les discours ordinaires. En définitive, comme le résumait laconiquement Hassan Abu Sa'ud, compositeur et président du Syndicat des métiers de la musique décédé en 2007, qui débuta lui-même sa carrière sur l'avenue : « Les musiciens de l'avenue Mohamed Ali jouent pour les petites gens » (entretien au siège du Syndicat, quartier de Bab al-Luq, 2004). [ant-art-245-body]

- [216] [...] Le second schéma d'aménagement (*Master plan*), rédigé en 1995, se fixait pour objectif l'année 2011. Ses auteurs considéraient que la réalisation du plan de 1975 n'avait pas été assez efficace, Chennai ayant connu une croissance démographique trop importante et les pôles ayant stagné par rapport à ce qui était escompté, même si la croissance du centre a été plus faible en pourcentage que celle de la périphérie. Malgré ce relatif échec constaté, il n'en demeurerait pas moins que les principales localisations de la croissance urbaine future étaient bel et bien en place.

En définitive, plus que la stratégie de planification urbaine conduite par les agences de développement de l'aire métropolitaine et la volonté récente des pouvoirs publics de vouloir faire des villes « des moteurs de la croissance », ce sont surtout les

politiques de libéralisation économique des années 90 qui sont à l'origine de la reconfiguration métropolitaine actuelle. [geo-art-79-body]

Nous relevons cinq occurrences (sur le total de 54) où *en définitive* prend sous sa portée un syntagme (**P_L**) : il peut s'agir d'un syntagme nominal [217] ou verbal [225]. En fait, la nature syntaxique de la séquence **q** est régie par celle de la séquence **p**. Quand la séquence **p** porte sur un syntagme nominal, la séquence **q** l'est également.

- [217] Aussi bien, cette dualité a été d'emblée dénoncée par les théoriciens : ils se sont toujours défendus de vanter le culte d'un instantanéisme punctiforme qui se porterait au hasard sur des thèmes hétérogènes. Ils n'ont cessé de déclarer, et de façon unanime, que leur intention n'est pas d'encourager la dispersion et, **en définitive**, l'instabilité et la fugacité mentales, mais d'honorer des préoccupations assez solides et profondes pour convaincre ceux qui les nourrissent d'endurer l'austérité liée à leur exploitation systématique. [sed-art-383-body]

c. Position syntaxique

En définitive se trouve en position initiale et introduit la séquence reformulatrice. Au niveau phrastique, il est placé en général en tête de phrase [218] ou en position postverbale [219]. *En définitive* peut se situer en position détachée pour introduire un paragraphe.

- [218] Cette expérience l'amène à questionner temps didactique et temps d'apprentissage, sans pour autant prendre en compte, dans l'analyse a p o s t e r i o r i, deux aspects ayant pu contribuer à la réussite du cycle: l'impact de l'activité golf sur la motivation des élèves et leur faible degré d'expertise en la matière. **En définitive**, un registre de valeurs se dessine, orientant largement la pratique professionnelle future. [sed-art-362-body]
- [219] Or ces open shrublands correspondent en réalité, pour une large part, soit aux milieux de tundra et de taïga des régions boréales canadiennes et russes, soit aux espaces désertiques australiens, autrement dit à des régions qui, tant par leur éloignement que par leurs conditions écologiques, ne présentent guère de potentialités de boisement. Les principales régions offrant en définitive des potentiels se situent dans les espaces semi-arides de Chine et des États-Unis et pour partie d'Australie. [geo-art-110-body]

Nous relevons uniquement deux occurrences de *en définitive* (sur le total de 54 occurrences) en fin de séquence [220][221]. Dans ces deux exemples, ce marqueur a une portée globale. Il porte sur un syntagme nominal dans [220] et sur un syntagme verbal dans [221].

- [220] Cependant, fondamentalement, la communication politique est plus vaste que la communication institutionnelle publique. Elle a trait aux relations sociales dans un espace de coopération et de conflit, au lien social en définitive. [infcom-art-601-body]
- [221] Ces prescriptions légales, parce qu'elles sont intégrées dans la pratique des organisations (Duran, 1992), sont un puissant facteur de renforcement du groupe des communicateurs publics. Il en résulte que leurs applications concrètes relèvent de l'expertise des professionnels de la communication, rendue plus nécessaire en raison des risques notamment de remise en cause des élections en cas de manquements. Plus

largement, l'activité communicationnelle est régie par une norme juridique impersonnelle qui la légitime **en définitive** (Laufer, Paradeise 1982). [infcom-art-601-body]

d. Valeurs sémantiques

Ce marqueur sert à marquer l'élément ou le point de vue considéré comme ultime dans une succession d'éléments ou points de vue [222][223].

[222] En réponse à notre questionnement, nous proposons de mettre en évidence trois formes de partage. Nous traiterons dans un premier temps du partage de l'expérience de visionnage, pour aborder ensuite le problème des conversations en ligne et des relations nouées sur la toile. Puis, après avoir relevé l'entremêlement de ces deux activités, nous terminerons par les questions d'échange et de redistribution de contenus sériels. Nous verrons que ces trois modalités de partage sous-tendent **en définitive** une tension entre les dimensions individuelle et collective de la consommation sérielle, que cette population de fans fait apparaître de manière saillante. [infcom-art-21-body]

[223] À l'inverse du français, la graphie des verbes anglais neutralise souvent des différences sonores (cf. les différentes prononciations de la terminaison -ed). De ce fait, à la différence du français, toute flexion écrite de l'anglais a un correspondant oral, ce qui facilite la scription. Toutefois, à une flexion écrite unique peuvent correspondre différentes prononciations. Comme en français, la transcription phonographique ne garantit donc pas la justesse morphographique. **En définitive, dans les deux langues, les apprenants doivent apprendre à se détacher de la forme orale pour produire la forme écrite, mais ils procèdent de manière différente.** [lin-art-225-body]

L'exemple [222] nous paraît assez bien structuré. Le locuteur y présente le plan de son travail et explique comment sont construites ses idées. La séquence **q** où apparaît *en définitive* correspond au dernier élément à traiter par le locuteur. En particulier, nous relevons un changement de perspective énonciative entre la séquence **p** et **q**. Si dans la séquence **p** sont présentés les objets d'études, la méthodologie de recherche, la séquence **q** introduite par *en définitive* exprime d'une manière explicite le point de vue de l'auteur en vue d'une préparation du lecteur au résultat.

En définitive joue le même rôle dans [223] quand la séquence reformulatrice insiste également sur le dernier point de vue faisant suite à une succession d'idées énoncée précédemment. Cet ultime point de vue qui porte sur l'enseignement et l'apprentissage du français et de l'anglais résulte d'une analyse antérieure sur la graphie et de la phonie de français et de l'anglais. À la différence des marqueurs d'énumération qui signalent seulement l'unité finale d'une série d'éléments comme *en dernier lieu, pour conclure, pour finir*, etc., le marqueur de reconsidération *en définitive* se caractérise par l'introduction du point de vue du locuteur.

Dans les exemples [224][225], l'emploi de *en définitive* a la même valeur. Dans [224], après avoir fait un bilan des recherches sur les *effets du redoublement*, la reformulation introduit le dernier point de vue du locuteur qui envisage en fait de les caractériser. Dans [225], *en définitive* porte sur un syntagme verbal et souligne l'opinion de l'auteur sur la position avantageuse des *chercheurs seniors* par rapport aux *jeunes chercheurs* dans cette étude. Cette formulation finale est annoncée au terme du processus d'analyse. Nous constatons que dans ces exemples, la séquence reformulatrice établie par *en définitive* coïncide avec la formulation attendue du locuteur.

- [224] Bref, il semble bien que l'on s'intéresse aux effets du redoublement depuis près d'un siècle, et ceci sans que la problématique soit considérée comme définitivement résolue. Or, sans pouvoir établir une comptabilité précise, il est manifeste que le nombre d'études sur le sujet est impressionnant : plus d'une centaine, estimons-nous. Plus surprenants encore, les résultats engrangés au fil de ces multiples efforts de recherche sont substantiellement cohérents. Rien n'y fait : l'incrédulité du monde de la pratique, mais aussi celle de certains décideurs politiques, semble bien irréductible. Pourtant, afin de répondre aux objections qui leur ont été et leur restent opposées, les chercheurs se sont efforcés d'affiner leur méthodologie de recherche. **En définitive, on peut considérer les recherches sur les effets du redoublement comme paradigmatiques en ce qui concerne les obstacles méthodologiques que tentent de contourner les chercheurs en éducation pour approcher une saisie fiable des effets d'une pratique d'enseignement.** [sed-art-482-body]
- [225] À plusieurs reprises, les chercheurs rencontrés se sont finalement réfugiés derrière leur statut pour justifier leur non-participation à ces forums publics. Ainsi, pour un chercheur, B, pour être en mesure de faire des commentaires, il faudrait se sentir à l'aise, ce qui est difficile lorsque l'on n'a pas la légitimité du rapporteur. Mais ce même chercheur évoque le fait que, depuis qu'il n'est plus jeune docteur, mais chargé de recherche CNRS, il se sent davantage en mesure d'intervenir, s'il y a des erreurs dans un manuscrit ou si une référence est incorrecte. Ce qu'il n'a pour l'instant jamais fait. Finalement, les jeunes chercheurs, par autocensure, laissent souvent les seuls chercheurs avérés et reconnus (c'est-à-dire bien souvent des chercheurs seniors) utiliser et, **en définitive s'approprier le forum de discussion mis en place par la revue, dans un phénomène de reproduction d'une hiérarchie institutionnelle déjà bien installée.** [infcom-art-432-body]

En définitive se distingue de *somme toute* par son fonctionnement syntaxique et sémantique [226][227]. Sur le plan sémantique, dans l'exemple [226], le locuteur explicite dans la séquence **p** les avantages d'un modèle avec un seul facteur commun pour l'enquête mensuelle d'activité dans l'entreprise. Il énonce dans la séquence reformulatrice son opinion sur l'application de la même méthode pour le travail d'interprétation. Dans l'exemple [227], la séquence **p** consiste en une description des travaux sur le village traditionnel d'un groupe de sociologie. Dans la reformulation introduite par *en définitive*, le locuteur insiste sur l'idée la plus intéressante qu'il retient des études antérieures pour son étude. L'utilisation du marqueur indique un changement d'angle de vue à partir duquel le locuteur peut introduire sa propre idée.

- [226] Pour ce qui concerne la présente étude, nous avons mis en évidence qu'un modèle avec un seul facteur commun peut décrire de façon satisfaisante les soldes d'opinion de l'enquête mensuelle d'activité dans l'industrie : ceci permet de justifier l'utilisation de ce facteur comme un indicateur résumé, sorte de climat général. De plus, dans ce cadre, l'apport supplémentaire des différentes questions peut être analysé à travers l'évolution des composantes spécifiques. En définitive, on peut espérer qu'un tel outil constitue une aide dans le travail d'interprétation délicat que doivent mener les conjoncturistes. [eco-art-188-conclusion]
- [227] [...] Même si les travaux du groupe de sociologie de Nanterre ont montré que ces caractéristiques de la société villageoise traditionnelle du XIX^e siècle (idéal type) s'observaient à une échelle plus étendue, celle du canton, et que les relations sociales des habitants s'inscrivaient dorénavant dans des espaces sociaux pluriels, ils ont aussi relevé que les municipalités étaient un lieu politique tout à fait important pour la régulation locale des relations sociales. En définitive, on retiendra habituellement de ces travaux la persistance d'une vie sociale intense à l'échelle des communes où l'interconnaissance constituerait un mode de repérage social permettant à chacun de connaître tout le monde et plusieurs aspects d'autrui. Dans un tel contexte de faible densité démographique et d'organisation sociale, il serait quasi impossible pour les habitants de ces communes de se rendre invisibles. [soc-art-503-body]

Dans ces exemples, *en définitive* porte sur une proposition. L'emploi de *somme toute* ne peut pas fonctionner dans ces exemples. Ce marqueur de reconsidération porte essentiellement sur un syntagme adjectival et sert à introduire le jugement de valeur du locuteur suite à une analyse d'éléments préalable, alors que la reformulation instaurée par *en définitive* insiste uniquement sur l'ultime point de vue de l'auteur. À la différence de *somme toute*, *en définitive* peut articuler des arguments et idées concordants et opposés.

La séquence **q** (*en définitive q*) révèle l'opinion de l'auteur d'une manière plus ou moins explicite avec une fréquente apparition des marques liées au locuteur comme *on*, *nous*, *notre*, etc. [226][227][228].

- [228] En revanche, lorsque l'argumentation est faible et que parallèlement le public est intéressé par le thème et/ou très attentif, l'expressivité du locuteur peut produire l'effet inverse. En effet, nos résultats révèlent que la source est perçue de façon moins sincère dans cette situation. Il est probable que la « théâtralisation » du communicateur soit perçue ici comme une manœuvre visant à doper une argumentation anémique. En définitive, notre travail indique à quel point l'expression faciale et les éléments prosodiques qui les accompagnent habituellement (silences, intonations) peuvent être un levier important de l'influence persuasive. Dès lors, l'intentionnalité stratégique de tels signaux ainsi que leur variabilité sur le plan interindividuel méritent d'être davantage explorées. Les professionnels de la communication verbale (hommes politiques, avocats, professeurs) usent-ils plus que d'autres de telles expressions ? Et si c'est le cas, sont-ils toujours sincères et spontanés ? [psy-art-260-body]

Quand la séquence **q** porte sur une proposition, le contexte de la séquence **p** est aussi assez large. Elle correspond à des idées et des points de vue différents du locuteur. Nous retrouvons dans le contexte gauche de *en définitive* les marqueurs d'enchaînement thématique (*dans un premier temps, puis, enfin*, etc.), les marqueurs de comparaison (*à l'inverse, comme*,

etc.). Comme *en définitive* marque le dernier point de vue du locuteur, le repérage des MD aide à repérer les idées essentielles développées dans le contexte gauche. Le paragraphe introduit par *en définitive* peut être le dernier d'une partie textuelle induisant ensuite un changement de partie.

Nous regroupons dans le Tableau 5.12 : Récapitulatif des propriétés syntaxiques et sémantiques de *en définitive* les propriétés essentielles de *en définitive*.

Paramètres	Propriétés
Position	Initiale. Au niveau phrastique, il peut se trouver en tête de phrase ou en position postverbale.
Portée	<p>Globale : <i>en définitive</i> porte en général sur une proposition, la séquence p est également de nature propositionnelle.</p> <p>Locale (très peu) : il porte sur des syntagmes nominal, adjectival ou verbal. La nature syntaxique de la séquence q est régie par la séquence p.</p> <div style="text-align: center;"> $\left(\begin{array}{c} \text{proposition} \\ \text{paragraphe} \end{array} \right)_{\mathbf{p}} \quad \text{En définitive} \quad \left(\begin{array}{c} \text{proposition} \end{array} \right)_{\mathbf{q}}$ $\left(\begin{array}{c} \text{Syntagme nominal} \\ \text{Syntagme verbal} \\ \text{Syntagme adjectival} \end{array} \right)_{\mathbf{p}} \quad \text{en définitive} \quad \left(\begin{array}{c} \text{Syntagme nominal} \\ \text{Syntagme verbal} \\ \text{Syntagme adjectival} \end{array} \right)_{\mathbf{q}} \Bigg)_{\text{phrase}}$ </div>
Valeurs sémantiques	<p>✓ <i>Valeur terminative</i> : <i>en définitive</i> souligne l'élément ou le point de vue considéré comme ultime du locuteur à la suite des éléments.</p> <p>➤ Quand il a une <u>portée globale</u> :</p> <p>[proposition₁+...+proposition_n/paragraphe]_p En définitive</p> <p> V.TERMINATIVE [proposition]_q</p> <p>➤ Quand il a une <u>portée locale</u> :</p> <p>[SN/SV/SA]_p en définitive V.TERMINATIVE [SN/SV/SA]_q</p>

Tableau 5.12 : Récapitulatif des propriétés syntaxiques et sémantiques de *en définitive*

Pour conclure, *en définitive* peut introduire une séquence à valeur terminative. À la différence des marqueurs d'enchaînement thématique (comme *en dernier lieu*, *en conclusion*, etc.) qui marquent la fin d'une étape ou d'autres marqueurs de reconsidération comme *somme toute* qui exprime l'opinion de l'auteur faisant suite à l'examen d'éléments antérieurs, *en définitive* révèle l'ultime point de vue de l'auteur tout en marquant un changement de perspective énonciative par rapport à son contexte gauche. La reformulation instaurée par *en définitive* correspond à la formulation sur laquelle le locuteur veut attirer l'attention du lecteur.

5.5.4. En fin de compte

a. Définition

En fin de compte est le seul marqueur qui a quasiment le même ratio dans notre corpus SHS et dans le corpus Emolex : 11 occurrences dans notre corpus (dont le ratio par 1 million de mots est de 0,04) contre 175 occurrences dans le corpus Emolex français général environ 40 M de mots (dont le ratio est de 0,04). Selon Roulet (1987), *en fin de compte* présente « la reformulation comme le résultat final de la prise en compte des éléments envisagés » (Roulet, 1987 : 130). La signification de la locution adverbiale est liée au sémantisme de ses composants *fin* et *compte*. Dans le TLFi, il a pour synonymes *au bout du compte*, *finalement*.

b. Portée syntaxique

En fin de compte porte en général sur une proposition (**P_G**) [232][233]. Dans l'exemple [229], la portée **q** correspond à la proposition subordonnée. Quelle que soit la position du marqueur, la séquence **p** (**p en fin de compte**) est toujours de nature propositionnelle.

- [229] La Coopool propose une combinaison singulière des activités politiques à la fois hors du parti et en ligne. Cette combinaison s'appuie sur une dissociation opérée entre la qualité de militant, devenue extrêmement plastique, et les débouchés du militantisme. S'appuyant sur l'individualisation des pratiques militantes, elle opère un brouillage virtuel des frontières partisans qui, **en fin de compte**, renforce l'efficacité des modes de sélection interne en les occultant au profit d'une représentation fédératrice d'un entre soi virtuel. [spo-art-340-body]

c. Position syntaxique

La position de *en fin de compte* est plus mobile par rapport aux autres marqueurs. Il se trouve souvent en position initiale quand il porte sur une proposition. Par ailleurs, il peut être placé en position médiane quand *en fin de compte* porte sur une proposition subordonnée [230]. En position initiale, il est toujours suivi d'une virgule [232][233].

- [230] Ce n'est donc qu'au terme d'un long mûrissement que le PS a reconnu statutairement ses militants en tant qu'individualités : des années 1990 où le PS renonce progressivement aux mandats, en passant par les adhérents à 20 euros et la Coopol aujourd'hui, l'individuation des pratiques militantes rejoint une individualisation des représentations partisans de l'engagement. En ce sens, Internet n'est qu'un prolongement technologique d'une évolution déjà ancienne et qui se donne à voir à travers les procédures de démocratisation des partis. L'apparente liberté des « coopains » socialistes nécessite, d'abord, de conjuguer la singularité militante au pluriel et de donner corps, ensuite, à l'injonction de « mise en responsabilité choisie » des militants politiques. La Coopol encourage ainsi un « empowerment militant » dont les représentants du parti sont, en fin de compte, les principaux maîtres d'œuvre et, surtout, les bénéficiaires. [spo-art-340-body]

d. Valeurs sémantiques

Par son sémantisme, *en fin de compte* peut introduire une séquence reformulatrice à valeur terminative. Celle-ci intervient au terme d'une analyse des éléments antérieurs et correspond au point de vue de l'auteur sur le résultat final [231][232].

- [231] Confrontés à une désaffection finalement souvent plus problématique en termes de légitimité et d'image qu'en termes de fonctionnement proprement dit, et au-delà des effets d'imitation, les partis politiques ont alors entrepris de « démocratiser » leur organisation, manière d'octroyer aux adhérents une place accrue dans les processus de décisions internes et de leur offrir des espaces de participation renforcés. Cette « démocratisation », dont les effets sont constatables à travers la mise en place de procédures délibératives et participatives souvent formelles, est perçue cependant, en fin de compte, comme encore insuffisante, comme l'attesterait la stagnation des adhésions. [spo-art-340-body]
- [232] Dans notre enquête sur les sports collectifs (Dugas, 2002), le temps consacré au jeu réel est considérablement réduit : en moyenne trois heures effectives pour dix séances d'éducation physique. Au regard de ces constats, nous pouvons nous demander si les objectifs poursuivis dans la spécificité de l'activité traitée sont atteints. De fait, une telle conduite didactique est-elle toujours pertinente sur des cycles aussi courts ? Par ailleurs, en situation de pratique « réelle » de tels ou tels sports, les professeurs d'EPS transforment souvent la nature de ces activités afin de répondre aux objectifs éducatifs et de s'adapter aux capacités des élèves. Dans ces conditions, ceux-ci jouent-ils réellement dans la pertinence du sport pratiqué ? En fin de compte, nous observons plutôt durant les phases de jeu « situation réelle » que le sport pratiqué à l'école se transforme en une « didactique » cours duquel certains traits de logique interne du sport pratiqué sont modifiés. [sed-art-52-body]

Dans l'exemple [231], le locuteur explicite dans la séquence **p** ce que représente la démocratisation mise en place par les partis politiques. La reformulation de *en fin de compte* donne lieu au jugement de valeur sur les effets de cette *démocratisation*. Dans l'exemple [232], le lecteur est conduit à suivre la progression de la pensée du locuteur en allant de ses constats à ses questionnements. La séquence **q** marque un changement de perspective énonciative par rapport à ce qui précède. En fait, le locuteur y indique son opinion sur le changement de *certaines traits du sport pratiqué dans les phases de jeu situation réelle*.

À première vue, *en fin de compte* peut ressembler à *somme toute*. Ces marqueurs servent à préciser le point de vue du locuteur à la suite d'une analyse de la totalité des éléments. Pourtant, tandis que *somme toute* instaure entre la première formulation et la séquence reformulatrice un rapport par généralisation, *en fin de compte* relie des idées à un jugement de valeur. La reformulation établie par ce marqueur est considérée comme la formulation attendue du locuteur et elle résulte d'un processus d'analyse préalable.

Dans les exemples [233][234], le lecteur suit la construction complexe de la réflexion du locuteur dont le point de vue est exprimé plus ou moins explicitement. Dans [233], la séquence reformulatrice introduite par *en fin de compte* apporte une estimation approximative sur la *régionalisation des schémas* à la suite de constats évoqués dans la séquence **p**. Dans l'exemple [234], la reformulation correspond à un jugement de valeur épistémique du locuteur : *en fin de compte* peut articuler des éléments opposés, précisément *les séries sérielles ou immobiles et les séries feuilletonesques ou évolutives*.

[233] On a bien un phénomène de régionalisation des schémas de circulation, privilégiant les liens de proximité sur ceux de longue distance, notamment depuis la disparition du bloc soviétique (1991). Cette mutation de grande ampleur se répercute aussi sur les liens intra-régionaux. En Asie du Nord-Est, le poids des pivots voisins décline dès la seconde période, puis celui des ports japonais (exemple, Yokohama, Osaka, Chiba, Nagoya, Kobe) se stabilise jusqu'aux années 1997-1999 tandis que les ports russes (ex. Nakhodka, Vladivostok) et sud-coréens (ex. Inch'eon, Pusan, Ulsan, P'ohang) dominant depuis 2000. Le port de Dalian en Chine passe ainsi devant Hong Kong et Shanghai de par sa proximité du port de Nampo. **En fin de compte**, l'Asie du Nord-Est concentre actuellement 86 % des escales des navires nord-coréens au niveau mondial, contre seulement 54 % il y a 20 ans, tandis que les ports sud-coréens à eux seuls atteignent 30 % du total contre seulement 5 % il y a 20 ans. [geo-art-255-body]

[234] Si les dichotomies proposées par les deux auteurs sont à prendre comme des cas limites idéaux, elles se rejoignent en partie dans les usages différenciés qu'en ont les sériphiles. Les séries sérielles ou immobiles se prêtent en effet davantage à des pratiques collectives (*Les Simpson*, *Friends*, *NCIS*...). Le respect de l'ordre des épisodes ainsi que leur suivi continu ne sont pas primordiaux, donnant lieu à des consommations potentiellement moins assidues, plus opportunistes et davantage appropriées au visionnage à plusieurs. Ces séries concernent en outre des genres plus légers, voire humoristiques à l'instar des sitcoms, appelant à davantage d'échanges et

de commentaires in situ. Au contraire, les séries feuilletonesques ou évolutives, en général dites « sérieuses » (dramas), invitent à une attention plus soutenue. Ainsi, les rendez-vous à deux ou plus autour d'un de ces feuilletons tiennent plus difficilement la distance et chacun préférera bien souvent en fin de compte avancer à son propre rythme. [infcom-art-21-body]

Les significations de *en fin de compte* et *en définitive* se rapprochent et elles sont souvent interchangeables. Pourtant, l'emploi de *en définitive* ne peut pas fonctionner dans ces derniers exemples. Certes, ces deux marqueurs permettent d'exprimer le point de vue du locuteur, mais dans [233][234] la séquence **q** ne peut être considérée comme le dernier point de vue du locuteur. Elle apporte ici son jugement au résultat final.

Paramètres	Propriétés
Position	Initiale quand <i>en fin de compte</i> porte sur une proposition. Il est également placé en tête de phrase. Médiane quand le marqueur porte sur un syntagme. Il se trouve en position postverbale.
Portée	Globale. La séquence p se compose de plusieurs propositions. Locale. La séquence q se trouve souvent dans la proposition subordonnée, et la séquence p est de nature propositionnelle. $\left(\begin{array}{c} \text{proposition} \\ \text{p} \end{array} \right) \text{ En fin de compte } \left(\begin{array}{c} \text{proposition} \\ \text{q} \end{array} \right)$ $\left(\begin{array}{c} \text{proposition} \\ \text{p} \end{array} \right) \left(\left(\begin{array}{c} \text{Début q} \\ \text{en fin de compte} \\ \text{Fin q} \end{array} \right) \right)_{\text{phrase}}$
Valeurs sémantiques	✓ <i>Valeur terminative</i> : <i>en fin de compte</i> introduit le jugement du locuteur sur le résultat final. ➤ Quand il a une <u>portée globale</u> : [proposition ₁ +...+proposition _n] _p en fin de compte [V.TERMINATIVE] [proposition] _q

	<p>➤ Quand il a une <u>portée locale</u> :</p> <p>[proposition₁+...+ proposition_n]_p en fin de compte_{V.TERMINATIVE} [syntagme verbal]_q</p>
--	---

Tableau 5.13 : Récapitulatif des propriétés syntaxiques et sémantiques de *en fin de compte*

En bref, la séquence reformulatrice établie par *en fin de compte* consiste en un jugement de l'auteur qui intervient à la suite d'une analyse d'un ensemble d'éléments. Ce MD se distingue de *somme toute* par son fonctionnement syntaxique et sémantique. D'une part, *somme toute* porte essentiellement sur un syntagme adjectival et sert à introduire une séquence à valeur résumante. D'autre part, *en fin de compte* met en relation des procès à un point de vue et il porte en général sur une proposition et un syntagme verbal. La reformulation correspond au résultat attendu du locuteur. Par rapport à *en définitive*, ces deux marqueurs sont souvent considérés comme synonymes l'un de l'autre. Cependant, alors qu'*en définitive* insiste sur l'ultime point de vue de l'auteur, l'introduction de la reformulation de *en fin de compte* donne lieu à un argument ou point de vue de valeur terminative du locuteur.

5.5.5. Au fond

a. Définition

Nous ne relevons que 12 occurrences de *au fond*. Dans notre corpus scientifique, *au fond* est utilisé en tant que locution adverbiale et marqueur de reformulation. Dans le corpus Emolex, cette locution est beaucoup plus fréquente : 2935 occurrences pour environ 40 M de mots, mais *au fond* apparaît beaucoup plus souvent en tant que locution prépositionnelle comme dans la structure « *au fond de* + N » (Ex : *au fond du trou/du puits/de ses yeux*, etc.). S'il est utilisé comme locution adverbiale, il désigne en général la dimension spatiale (Ex : *plus froide au fond, elle est trop loin au fond*), ou la nature intrinsèque d'une chose (Ex : *Ils le regardent au fond*, etc.). Dans le cadre de notre étude, nous nous intéressons uniquement aux emplois discursifs de ce marqueur. Roulet (1987) considère la reformulation introduite par *au fond* « comme un examen approfondi du problème » (*ibid.* : 126). La signification du marqueur s'explique en quelque sorte par son composant substantif *fond* qui indique la

profondeur de l'analyse du contexte gauche de *au fond*. Dans le TLFi, *au fond* prend pour synonymes *tout bien considéré, en dernière analyse*.

b. Portée syntaxique

Au fond ne prend sous sa portée qu'une proposition (**P_G**) [235]. La séquence **p** est aussi de nature propositionnelle.

- [235] À la manière d'Aristote, pour qui l'être se déclinait en plusieurs sens, la substance cartésienne se décline donc au moins en deux. À vouloir résumer le spinozisme à une opération de technique scolastique, on pourrait dire que celle-ci ne fait **au fond** que clarifier cette ambivalence, en démasquant les ambiguïtés ou les impensés sur lesquels en dernière instance elle repose. Après quoi, pourront être avancées des thèses inouïes pour l'époque. [ant-art-627-body]

c. Position syntaxique

Au fond se trouve toujours en position initiale pour introduire la séquence reformulatrice [236]. Au niveau phrastique, il est placé souvent en tête de phrase ou bien en position postverbale [237]. En tête de phrase, il est suivi d'une virgule [236][238].

- [236] Ce serait donc à partir de ce relatif et décevant constat d'échec - et dans le but louable de la sauver - que les propositions d'une nouvelle manière d'envisager l'anthropologie ont été offertes sur le marché de la pensée scientifique. **Au fond**, c'est ce déficit ontologique que les premiers anthropologues cognitivistes français ont dénoncé et c'est la brèche dans laquelle ils se sont précipités : les notions employées par la discipline ne décriraient rien dont on puisse affirmer l'existence (Sperber 1996 : 28, 31). Tout se passerait comme si on avait affaire à des mots qui existent bien (« mariage, sacrifice, chefferie ») et dont on est capable de montrer le sens dans des définitions, mais qui ne « nous disent pas si les mariages, les sacrifices ou les chefferies font partie des choses qui meublent le monde » (Ibid. : 31). [ant-art-41-body]
- [237] [...] le nazisme, selon le sociologue allemand et chroniqueur de son temps, prétendait être édificateur d'un monde, orchestrait le concert des pelles et des truelles dans la symphonie des grands travaux, mais ne visait **au fond** qu'à incendier le monde dans une catastrophe nihiliste, une jouissance du néant qui considérerait la destruction comme une fin en soi. [his-art-495-body]

d. Valeurs sémantiques

Dans ce travail, nous analysons *au fond* comme un marqueur de reformulation qui sert à exprimer une reconsidération de l'auteur. En fait, par son sémantisme, *au fond* introduit le point de vue de l'auteur qui correspond à une analyse en profondeur de la première formulation. La reformulation par ce marqueur a donc une *valeur conclusive* [238][239].

- [238] Le signe est en quelque sorte impuissant, cognitivement parlant, à « représenter l'objet et à dire quelque chose de lui » (Peirce 1931-1935 : II, 231), car il ne représente qu'une connaissance dérivée des objets qui ont été connus une première fois par l'observation « collatérale ». Si l'on dit à quelqu'un que « Napoléon était un être léthargique » et que cette personne n'a jamais auparavant entendu parler de Napoléon, il ne parviendra pas à l'interpréter correctement dans la mesure où, pour lui, n'importe qui pourrait s'appeler Napoléon (Chauviré 1995 : 138). Au fond, tout se passe comme si on montrait à quelqu'un « un portrait de Leopardi, avec “Leopardi” écrit au-dessous » (Peirce 1958 : VIII, 183). Le portrait « transmet une information à une personne qui sait qui est Leopardi, et, aux autres, elle dit seulement que quelque chose du nom de Leopardi ressemblait à ceci » (Ibid.). [ant-art-41-body]
- [239] À travers la valorisation de l'emploi étudiant se joue une redéfinition du travail académique et des modes d'accès aux diplômes, dans laquelle la place accordée aux savoirs proprement universitaires tend à se réduire. En effet, les discours célébrant les vertus du travail en cours d'études participent d'une vision apologétique de la « professionnalisation » de l'enseignement, vision selon laquelle le système éducatif doit non seulement délivrer des connaissances, mais aussi favoriser le développement de « compétences ». L'entreprise est conçue comme le lieu d'acquisition par excellence de ces manières d'être et l'« expérience professionnelle », comme la meilleure voie possible pour leur mise à l'épreuve et leur perfectionnement, bref, comme l'expérience la plus « formatrice ». Au fond, il s'agit d'introduire dans un univers relativement autonome l'université des exigences et des catégories de classement issues du monde économique, notamment par la promotion des « savoir-être » valorisés dans les entreprises. [soc-art-118-body]

Dans l'exemple [238], la séquence reformulatrice peut être considérée comme une précision ou un approfondissement de ce qui est indiqué dans la séquence **p**. Cette reformulation sert à mieux illustrer l'idée précédemment évoquée sur *l'impuissance du signe*. Si *en somme* instaure entre la première formulation et la séquence reformulatrice un rapport par induction, dans cet exemple, le rapport établi par *au fond* semble être inversé, il fonctionne par déduction.

L'emploi de *en somme* dans l'exemple [239] ne fonctionne pas, car le locuteur n'envisage pas de faire un bilan ni de tirer des conclusions, mais de mieux clarifier la séquence **p**. *Au fond* se différencie également des marqueurs servant à clarifier (*c'est-à-dire, à savoir, en d'autres termes, autrement dit*). La reformulation effectuée par ces derniers correspond en général à une paraphrase, tandis qu'*au fond* marque le changement de perspective énonciative et exprime le point de vue du locuteur. La séquence reformulatrice introduite par *au fond* correspond à une analyse approfondie des arguments dans la première formulation. Dans l'exemple [239], nous relevons une seule idée principale dans la séquence **p** où le locuteur propose des pistes pour une valorisation de l'emploi des étudiants y compris la *professionnalisation de l'enseignement* et de possibles expériences que les étudiants peuvent acquérir à l'entreprise. La séquence reformulatrice explique plus précisément

comment s'effectue la valorisation de l'emploi des étudiants. Il s'agit en fait de l'intégration de quelques compétences de l'entreprise au milieu de l'université.

La reformulation de *au fond* sous-entend un processus d'analyse préalable du locuteur [240]. Nous remarquons une seule idée principale dans la séquence **p**. Dans l'exemple [240], la séquence **p** traite le rapport entre le statut et le prix des biens. La séquence **q** consiste en un examen des arguments dans la séquence **p**. Dans l'exemple [240], l'emploi de *en définitive* est impossible, car le locuteur ne cherche pas à donner son ultime point de vue. Dans l'exemple [241], la séquence reformulatrice où intervient *au fond* a pour l'objectif de compléter les idées manquantes dans la séquence **p**. En fait, elle apporte plus de précisions sur les indemnités des élus nationaux avant *les lois sur le financement de la vie politique de 1988-1990*.

[240] *Le statut tend à faire augmenter le prix de vente.* Un statut élevé est un signal de qualité. Or, nous dit-il, « les gens sont disposés à payer plus cher des biens dont ils croient que la qualité est meilleure » pour trois raisons : la préférence pour la qualité, la consommation ostentatoire, et les coûts élevés de recherche d'information sur la qualité des biens similaires. Il en conclut que les prix des biens de statut plus prestigieux tendent à être plus élevés, toutes choses égales par ailleurs. En quoi consiste, au fond, cette explication du prix ? Tout d'abord en une proposition que l'on peut considérer comme sociologique, sur l'impact du statut sur la demande : les consommateurs sont disposés à payer plus cher des biens dont le statut est élevé que d'autres biens de qualité effective équivalente, mais de moindre statut. Dire que les consommateurs sont « désireux de payer plus cher » revient exactement à dire que la demande est plus élevée. [soc-art-552-body]

[241] Mais ce processus implique des coûts que les élus hésitent à exposer clairement. Ainsi, des indemnités au lieu de salaires comme pour mieux dénier l'idée d'une contrepartie financière à l'engagement politique sont introduites très progressivement tant à l'égard des élus que de leurs collaborateurs. Les élus locaux et leurs auxiliaires sont touchés plus tard encore que les parlementaires par ce mouvement qui ne s'est accéléré qu'après les lois sur le financement de la vie politique française de 1988-1990. Et au fond, avant 1980, les élus nationaux devaient se débrouiller pour être assistés dans leur travail : pour payer leurs collaborateurs, il leur fallait soit trouver des sources occultes de financement, soit les payer sur leurs fonds propres. [spo-art-323-body]

Le Tableau 5.14 : Récapitulatif des propriétés syntaxiques et sémantiques de *au fond* récapitule les propriétés syntaxique et sémantique de *au fond*.

Paramètres	Propriétés
Position	Initiale. Au niveau phrastique, il est en général placé en tête de phrase. Médiane. Il peut se trouver en position post-verbale.

Portée	<p>Globale. Les séquences p et q sont de nature propositionnelle.</p> $\left(\begin{array}{c} \text{proposition} \\ \text{paragraphe} \end{array} \right)_p \quad \text{Au fond} \quad \left(\begin{array}{c} \text{proposition} \end{array} \right)_q$ $\left(\begin{array}{c} \text{proposition} \end{array} \right)_p \left(\left(\begin{array}{c} \text{Début } q \end{array} \right) \quad \text{au fond} \quad \left(\begin{array}{c} \text{Fin } q \end{array} \right) \right)_{\text{phrase}}$
Valeurs sémantiques	<p>✓ <i>Valeur conclusive</i> : la séquence reformulatrice correspond à une analyse en profondeur de la première formulation.</p> <p>[proposition/paragraphe]_p Au fond V.CONCLUSIVE [phrase]_q</p>

Tableau 5.14 : Récapitulatif des propriétés syntaxiques et sémantiques de *au fond*

Pour conclure, la séquence reformulatrice introduite par *au fond* souligne une analyse en profondeur des éléments de la première formulation. Elle clarifie le point de vue de l'auteur. *Au fond* se distingue des marqueurs servant à clarifier par son fonctionnement sémantique. La reformulation introduite par ces marqueurs consiste en une paraphrase, tandis que la séquence reformulatrice introduite par *au fond* indique l'opinion du locuteur et insiste sur une étude approfondie des idées et arguments antérieurs.

Dans le Tableau 5.15 : Comparaison des caractéristiques de *au total*, *somme toute*, *en définitive*, *en fin de compte* et *au fond*, figurent les propriétés syntaxiques et sémantiques essentielles des MD servant à récapituler afin de faire ressortir les spécificités d'emploi de chacun.

MD	Propriétés
Au total	<ul style="list-style-type: none"> - Il peut se trouver en position détachée au début du paragraphe. - Valeur épistémique : <i>au total</i> met en relation des arguments à un point de vue/argument/jugement. Il indique le cheminement d'analyse/de raisonnement (induction).

	<ul style="list-style-type: none"> - <i>Au total</i> permet d'articuler des idées contradictoires et concordantes. - La séquence p est composée de plusieurs idées et elle est marquée par l'emploi de différents MD. Il est important d'introduire ce phénomène dans l'enseignement/apprentissage des MD, car il permet à l'étudiant de relever les idées principales et les structures discursives.
Somme toute	<ul style="list-style-type: none"> - La séquence p renvoie au contexte antérieur pour donner le jugement du locuteur. - Il porte sur un syntagme adjectival. - La séquence p consiste en une seule idée ou des éléments concordants. ➤ Il peut être remplacé par <i>en somme</i> quand <i>somme toute</i> porte sur une proposition.
En définitive	<ul style="list-style-type: none"> - La séquence p est composée de plusieurs idées ou arguments. - Il marque un ultime point de vue. - En comparant avec <i>en définitive</i>, <i>somme toute</i> le distingue en portant essentiellement sur un syntagme adjectival. - <i>En définitive</i> et <i>en somme</i> se différencient en ce que le deuxième n'indique pas le jugement du locuteur.
En fin de compte	<ul style="list-style-type: none"> - La séquence p met en relation les procès et le jugement du locuteur sur le résultat final. - La différence entre <i>somme toute</i> et <i>en fin de compte</i> consiste en ce que ce dernier souligne l'analyse de la totalité des éléments. - A la différence de <i>en fin de compte</i>, <i>en définitive</i> insiste sur le jugement considéré comme ultime du locuteur. - En comparaison avec <i>au total</i>, <i>en fin de compte</i> insiste sur le jugement du résultat final. - Par rapport à <i>en fin de compte</i>, <i>au fond</i> sert à souligner une analyse en profondeur.

Au fond	<ul style="list-style-type: none"> - La séquence p correspond souvent à une seule idée. - Le marqueur introduit une analyse en profondeur de la première formulation. L'utilisation de ce marqueur sous-entend un processus d'analyse préalable du locuteur.
----------------	---

Tableau 5.15 : Comparaison des caractéristiques de *au total*, *somme toute*, *en définitive*, *en fin de compte* et *au fond*

Conclusion

Pour conclure ce chapitre, nous confirmons un lien étroit entre le fonctionnement syntaxique et sémantique des MD. En fait, l'association des critères syntaxiques (la portée syntaxique et la position) contribue au repérage des valeurs sémantiques. En outre, le travail d'analyse que nous avons effectué montre l'intérêt d'une telle étude sur corpus, car celle-ci permet de relever les propriétés syntaxiques et sémantiques des MD de reformulation, en particulier la comparaison des MD pour l'objectif linguistique et didactique que nous avons précisé. Elle sert à mettre en avant les spécificités les plus saillantes dans l'utilisation de chaque MD. Les tests de compatibilité nous permettent de justifier si un MD est remplaçable par un autre et dans quel contexte.

La reformulation établie par les MD peut être plus ou moins subjective. Si la reformulation établie par les marqueurs de clarification et de récapitulation est relativement neutre, les marqueurs de reconsidération se distinguent par l'empreinte subjective du locuteur dans la séquence reformulatrice. La reformulation dans la séquence **q** correspond au jugement ou au point de vue considéré comme ultime du locuteur (*en définitive*) à la suite d'une analyse préalable plus ou moins longue (*au total*, *au fond*) et au terme d'un processus de raisonnement plus ou moins complexe (*en fin de compte*).

La taille du contexte gauche du MD est très variable d'un MD à l'autre. Les marqueurs servant à clarifier (*c'est-à-dire*, *à savoir*, *autrement dit*, *en d'autres termes*) ont une séquence **p** relativement courte qui se limite au niveau d'une phrase. La taille de la séquence **p** des marqueurs servant à récapituler et à reconsidérer est relativement grande, surtout pour les marqueurs récapitulatifs (*en somme*, *en résumé*) qui ont un contexte gauche parfois assez volumineux. En effet, ces marqueurs permettent de reformuler plusieurs idées pour faire ressortir les idées principales. Nous relevons également des MD d'autres catégories (MD

d'enchaînement thématique, de comparaison, de topicalisation, etc.) dans la séquence **p** de ces marqueurs. Ce phénomène est très important pour le développement des stratégies de lecture que nous développerons dans le chapitre suivant, car il sert à tracer les idées essentielles dans la séquence **p**.

Enfin, des descriptions linguistiques très détaillées des marqueurs de reformulation effectuées dans ce chapitre nécessitent d'être exploitées pour des activités didactiques. Il s'agit de l'objectif principal du chapitre 6 où nous proposerons des applications didactiques en lien avec cette étude linguistique sur corpus. Ces descriptions et la typologie des séquences lexicalisées à fonction discursive conçue dans le chapitre 4 nous serviront de base pour les réflexions didactiques ainsi que les propositions didactiques que nous allons aborder dans le chapitre 6. Elles nous apportent à la fois des données riches permettant de proposer une méthodologie adaptable à l'enseignement/apprentissage des MD, mais aussi d'autres unités lexicales.

PARTIE 3 :

RÉFLEXIONS DIDACTIQUES

Chapitre 6.

Propositions didactiques pour l'enseignement/apprentissage de séquences lexicalisées à fonction discursive

Nous avons procédé dans le chapitre 4 et dans le chapitre 5 à une classification des séquences lexicalisées à fonction discursive et à une analyse syntactico-sémantique des marqueurs de reformulation. Ces descriptions linguistiques s'avèrent indispensables pour mieux comprendre leur fonctionnement syntaxique et sémantique et révéler comment elles participent à la structuration du texte. Cependant, ces descriptions ne sont pas directement accessibles aux enseignants de FLE qui ne sont pas des linguistes et n'ont pas l'habitude d'accéder à de longues descriptions linguistiques. En particulier, pour enseigner ces séquences lexicalisées à fonction discursive aux apprenants, il est nécessaire d'effectuer une transposition didactique de ces connaissances.

Dans ce chapitre, nous souhaitons proposer aux enseignants de FLE différentes approches pour intégrer les MD dans les contenus d'enseignement. En effet, les réflexions didactiques que nous allons présenter dans cette partie n'auraient pas pu se faire sans la description préalable des traits syntaxiques et sémantiques des MD dans le chapitre 5. Les enseignants de FLE qui sont intéressés par nos suggestions didactiques pourront approfondir ultérieurement leurs connaissances en consultant les analyses linguistiques. Ce chapitre s'organise de la manière suivante : premièrement, nous effectuerons une analyse du corpus d'apprenants vietnamophones afin d'avoir un exemple des difficultés et des erreurs des apprenants utilisant les marqueurs discursifs. Deuxièmement, nous mettrons l'accent sur l'importance de l'utilisation du corpus au service de l'enseignement et de l'apprentissage des langues et nous développerons ainsi les stratégies d'enseignement et définirons le rôle de

l'enseignant dans la transposition didactique. Enfin, nous proposerons différentes réflexions didactiques qui traiteront essentiellement de l'enseignement des MD tout en mettant en place quelques fiches-types pédagogiques.

6.1. Corpus de mémoires d'apprenants

Précisons de prime abord que le corpus de mémoires d'apprenants vietnamophones n'a pas été constitué dans le but de le mettre en comparaison avec notre corpus de travail (corpus d'articles scientifiques en SHS). Il s'agit en effet de deux genres différents, qui ne sont pas comparables. Notre corpus de travail composé d'articles scientifiques en SHS et rédigés par des locuteurs natifs nous sert à relever les régularités de fonctionnement syntaxique et sémantique des marqueurs discursifs. Ces articles ont fait l'objet d'une évaluation par un comité de relecteurs dans le même domaine. Nous considérons ces articles comme le modèle vers lequel les étudiants peuvent se diriger afin de mieux répondre aux exigences des écrits scientifiques aux plans structurel, syntaxique et sémantique⁴⁶.

Le corpus d'apprenants pour sa part nous permettra de faire un diagnostic de l'utilisation des MD chez des étudiants de niveau avancé pour en révéler les lacunes et les erreurs, ce qui permettra ensuite de proposer des activités qui sont susceptibles de répondre au mieux aux besoins spécifiques de ces étudiants dans la rédaction de leur travail.

6.1.1. Élaboration

Notre corpus de mémoires d'étudiants vietnamiens est de 305 568 mots. Ce corpus se compose, en règle générale, de mémoires de fin d'études universitaires (au Vietnam la licence se fait en quatre ans) qui relèvent des disciplines suivantes : économie, géographie, psychologie et linguistique⁴⁷. Il est à noter que ces mémoires ont été rédigés en français et corrigés par des enseignants de français d'origine vietnamienne ; à l'exception des deux mémoires de master en psychologie qui ont été rédigés dans le cadre du master de coopération entre une université vietnamienne et une université française et évalués par des enseignants français (Tableau 6.1).

⁴⁶ Cf. Chapitre 1 écrit scientifiques.

⁴⁷ Nous tenons à remercier tous les étudiants de nous avoir gentiment accordé le droit d'utiliser leur mémoire pour cette recherche en thèse.

Disciplines	Nombre de mémoires
Économie	5
Géographie	9
Linguistique	6
Psychologie	2
Total	22

Tableau 6.1 : Corpus de mémoires d'apprenants vietnamophones

Toutes les parties d'annexes ont été enlevées, par exemple, la page de garde, la table des matières, les références bibliographiques, ainsi que les tableaux et les figures. Ce processus de nettoyage s'est avéré fastidieux, mais nécessaire pour le traitement de ce corpus d'apprenants. En effet, nous nous sommes intéressée uniquement à la façon dont les étudiants structurent leur écrit dans la partie introduction, développement et conclusion. Par ailleurs, comme nous envisageons d'utiliser le logiciel Antconc compatible uniquement avec les fichiers en format texte (format *.txt*), les images, les figures et les tableaux doivent être supprimés du fichier.

L'objectif de la constitution de ce corpus d'apprenants est d'effectuer une étude d'ordre quantitatif et qualitatif. Nous souhaitons principalement : 1/ relever la fréquence des organisateurs textuels dans leurs écrits, et 2/ analyser comment les étudiants les utilisent. Le logiciel Antconc est un logiciel simple à manipuler qui répond bien à nos besoins. Cet outil permet d'effectuer des recherches par mots clés, des calculs de fréquences, d'accéder aux concordances, ainsi que de travailler sur plusieurs textes simultanément. Développé par Laurence Anthony, ce logiciel est téléchargeable à l'adresse suivante : <http://www.antlab.sci.waseda.ac.jp/>.

6.1.2. Analyse

Afin de faciliter l'analyse et de respecter l'anonymat des mémoires des apprenants, les mémoires ont été nommés par leur discipline et numérotés. La liste détaillée des sujets de mémoires de recherche correspondant à leur énumération se trouve dans l'*Annexe 3 : Corpus de mémoires d'apprenants vietnamophones*. La Figure 6-1 correspond à la copie d'écran d'une recherche des occurrences de *autrement dit*. À gauche, dans la colonne **Corpus Files**, s'affiche la liste des mémoires des apprenants. Au centre, se trouve la concordance de *autrement dit*, marqueur qui apparaît 26 fois dans notre corpus.

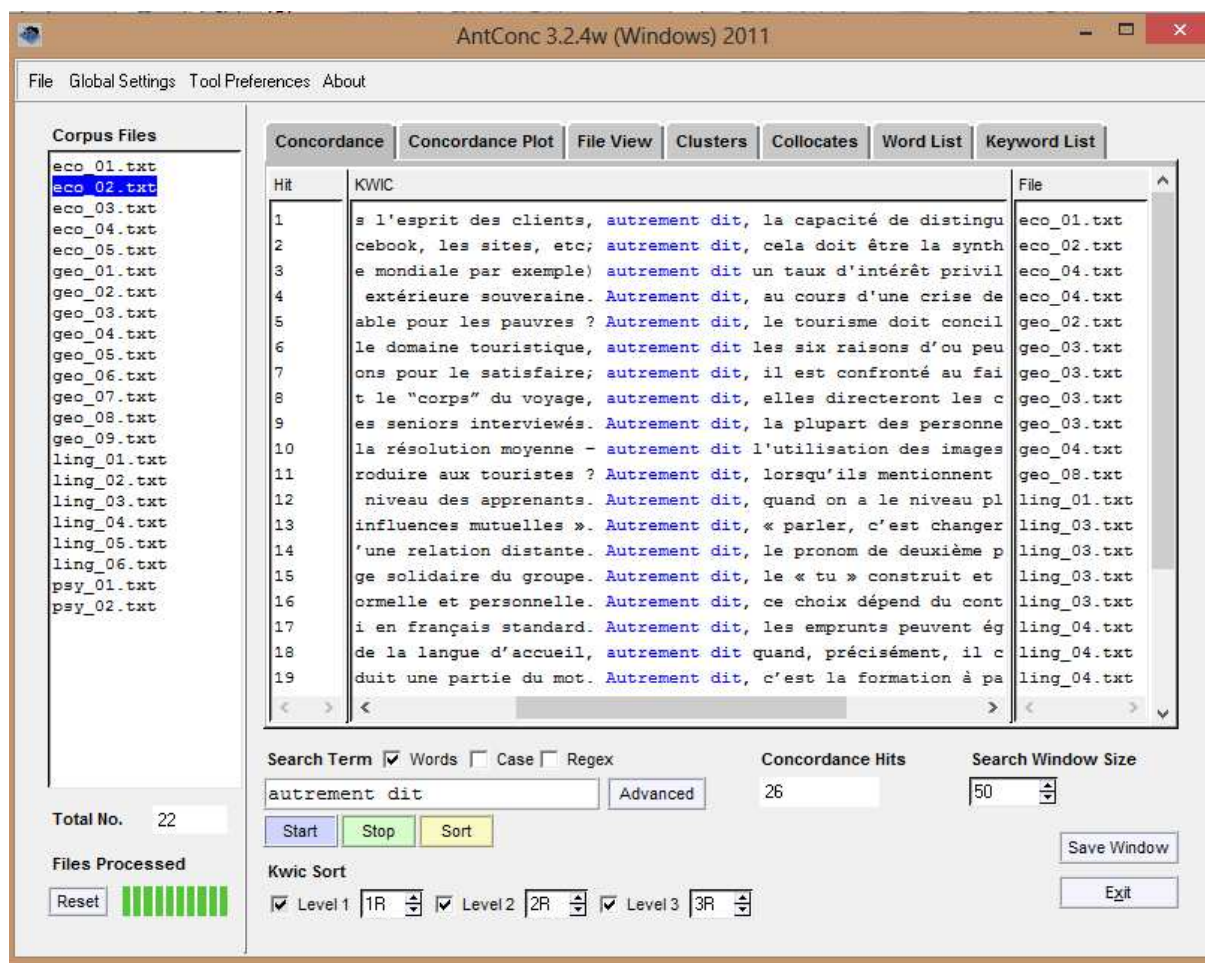


Figure 6-1 : Concordance de *autrement dit* dans le corpus d'apprenants

Antconc offre la possibilité de naviguer entre chaque occurrence retrouvée et son contexte large. Nous pouvons donc analyser comment les étudiants utilisent les organisateurs dans chaque cas précis. La Figure 6-2 représente le contexte large d'une occurrence de *autrement dit*.

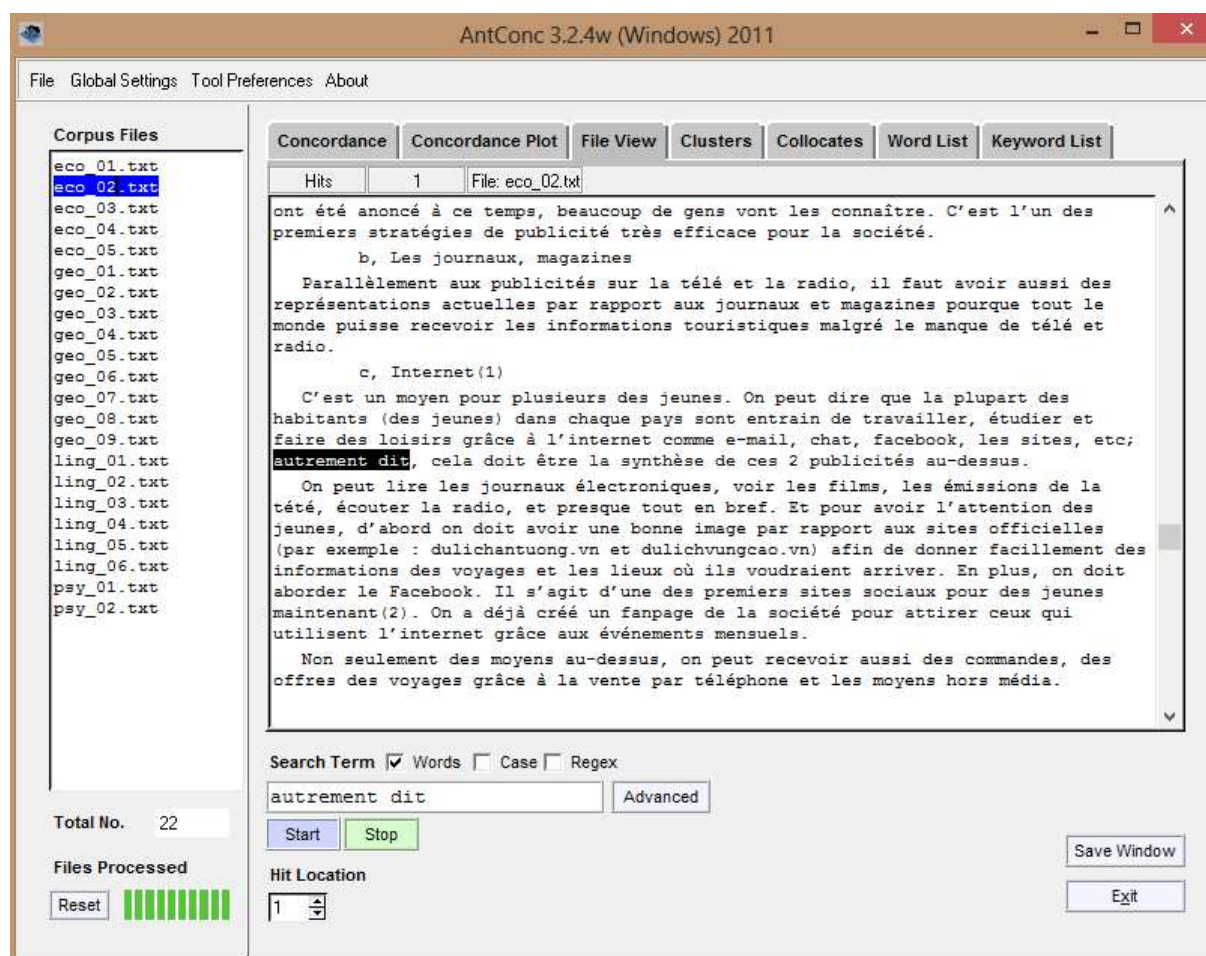


Figure 6-2 : Navigation entre les occurrences et le contexte large

L'analyse du corpus des mémoires des apprenants nous permet de faire quelques remarques concernant l'utilisation des organisateurs textuels chez les étudiants. Nous noterons d'abord le faible emploi d'organiseurs textuels, des MD dans leur écrit ; puis l'utilisation inappropriée des quelques marqueurs présents.

Au plan quantitatif, les marqueurs discursifs sont très peu fréquents dans les mémoires des étudiants. Les autres marqueurs monolexicaux ou bien les *expressions logiques* (appellation récurrente dans de nombreuses méthodes de français) sont beaucoup plus fréquents que les marqueurs polylexicaux. Par exemple, nous enregistrons 316 occurrences de *alors*, 255 occurrences de *donc*, 164 occurrences de *pourtant*, 101 occurrences de *cependant*, etc. contre 22 occurrences de *d'abord*, 10 occurrences de *en premier lieu*, 12 occurrences de *à savoir*, etc. Si nous comparons l'utilisation des marqueurs polylexicaux dans le corpus de mémoires d'apprenants et dans notre corpus SHS, nous relevons 144 occurrences de *d'abord*, 88 occurrences de *en premier lieu*, 139 occurrences de *à savoir*. La fréquence relativement

faible et l'absence de l'utilisation des marqueurs discursifs s'expliquent, selon nous, par le manque de présentation des organisateurs textuels dans les méthodes, tandis que les expressions monolexicales sont davantage enseignées à tous les niveaux. Cependant, il est à noter que quelques marqueurs polylexicaux sont plus fréquents que d'autres dans ce corpus d'apprenants : il s'agit des cas de *de plus* (140 occurrences), *en outre* (77 occurrences), *parce que* (63 occurrences), *tandis que* (75 occurrences), *par contre* (42 occurrences), etc. Nous supposons que le taux relativement élevé de ces marqueurs est lié au fait qu'ils sont davantage abordés dans les méthodes de français que d'autres.

Il manque, nous semble-t-il, un enseignement systématique de ces marqueurs polylexicaux dans les méthodes (Cf. Chapitre 2) et par conséquent, de la part des enseignants. Les apprenants ignorent le rôle de ces marqueurs discursifs dans la structuration du discours, soit parce qu'ils ne se sentent pas à l'aise en utilisant les organisateurs textuels, soit parce qu'ils ne savent pas comment les utiliser. Nous pouvons donc affirmer l'importance de prendre en considération ces marqueurs discursifs comme objet d'enseignement en tant que tel afin d'aider les apprenants à comprendre leur fonctionnement pour ensuite les réinvestir dans leurs écrits. Sont-ils aussi présents dans les écrits scientifiques vietnamiens ? Existerait-il une corrélation entre la traduction de ce type d'écrit par les étudiants ? Des entretiens avec les apprenants et les enseignants nous semblent intéressants à envisager. Ils serviraient d'un côté à évaluer les habitudes d'utilisation des marqueurs discursifs chez les apprenants et à révéler les raisons du faible emploi d'utilisation de ces organisateurs. D'un autre côté, nous pourrions adapter l'enseignement afin de développer différentes stratégies d'apprentissage chez les apprenants.

Nous recensons dans le Tableau 6.2 la fréquence et le ratio des marqueurs de reformulation dans notre corpus SHS et dans le corpus de mémoires d'apprenants. Ce tableau sert uniquement à mettre en évidence la différence d'utilisation des marqueurs dans les deux corpus.

Marqueurs de reformulation	Fréquence dans le corpus SHS		Fréquence dans le corpus d'apprenants	
	Nombre d'occurrences	Ratio (/10 000)	Nombre d'occurrences	Ratio (/10 000)
C'est-à-dire	704	2,63	33	1,08
Autrement dit	197	0,74	26	0,85

À savoir	139	0,52	12	0,39
En d'autres termes	119	0,44	2	0,07
En somme	57	0,21	0	0
En résumé	21	0,08	4	0,13
En un mot	14	0,05	1	0,03
Au total	49	0,20	2	0,07
Somme toute	13	0,18	0	0
En définitive	54	0,05	1	0,03
En fin de compte	11	0,04	0	0
Au fond	12	0,04	0	0

Tableau 6.2 : Fréquence des marqueurs de reformulation dans le corpus d'apprenants vietnamophones

Dans le Tableau 6.2, nous relevons très peu de marqueurs qui servent à récapituler, *en résumé* est le seul marqueur de ce sous-groupe utilisé par les étudiants, tandis qu'*en somme* qui est très fréquent dans notre corpus SHS (57 occurrences), n'est jamais utilisé par les apprenants. Nous remarquons également un manque de marqueurs de reconsidération (*au total*, *en définitive*, etc.) qui ne sont quasiment pas utilisés par les étudiants. Cette absence nous semble peu surprenante, car ces marqueurs proposent une reformulation tout en introduisant le point de vue de l'auteur. Une des explications principales est sûrement un manque d'enseignement de ces marqueurs dans les programmes. C'est cet aspect que nous souhaitons analyser et apporter notre réflexion didactique pour un enseignement plus systématique. En outre, comme l'a mentionné Mroué (2014), l'absence d'utilisation de ce genre de marqueurs n'est pas relative à une culture, mais il s'agit d'un phénomène fréquent chez les jeunes « chercheurs » et/ou les étudiants en devenir. Cela reste encore à prouver, mais selon Mroué, les étudiants non natifs qui viennent faire leur Master en France n'ont souvent jamais rédigé de mémoire dans leur pays et ne savent, a priori, pas mieux le faire dans leur langue maternelle.

Par ailleurs, nous enregistrons des cas où le marqueur polylexical est utilisé de manière inappropriée. Nous concentrons notre étude ici sur l'utilisation des marqueurs de reformulation chez les étudiants en comparaison avec les descriptions linguistiques de ces MD abordées dans le chapitre 5. Dans l'exemple suivant, l'utilisation de *autrement dit* n'est

pas possible dans le contexte. Le lecteur aura donc du mal à comprendre le rapport entre la phrase qui précède *autrement dit* et celle qui le suit.

- [1'] Chaque ethnie possède sa propre langue, sa propre culture, faisant partie intégrante de la culture, riche en genres et en couleurs. Ce sont des conditions nécessaires pour développer du tourisme à ce parc. Mais la question est de savoir, comment assurer le développement du tourisme durable pour les pauvres ? **Autrement dit**, le tourisme doit concilier 3 aspects : économie, société et environnement. [geo_02]

Comme nous l'avons vu, *autrement dit* sert à introduire une reformulation qui correspond à une autre manière de dire la proposition précédente. Pourtant, dans cet exemple, la phrase introduite par *autrement dit* ne clarifie pas la question précédemment posée dans *la question est de savoir... ?*, mais plutôt à développer une idée. L'utilisation de *autrement dit* est inadéquate. Ce marqueur peut être remplacé selon nous par un autre comme *en fin de compte* qui est un marqueur de reformulation qui sert à reconsidérer.

Nous relevons d'autres cas où l'apparition du marqueur discursif n'est pas convaincante. Dans l'exemple [2'], l'apparition de *à savoir* nous paraît ambiguë.

- [2'] Ce résultat nous sert plus tard à confirmer ou infirmer l'idée que malgré les expériences différentes en français, certaines stratégies d'écoute sont applicables pour tous les apprenants. Nous sommes ensuite entrés en la compréhension orale avec des questions **à savoir** les attitudes des étudiants envers cette compétence. [ling_01]

Le problème ne se situe pas au plan syntaxique, mais plutôt au plan sémantique. Comme *à savoir* peut introduire une énumération ou un élément spécifique, dans cet exemple, nous ne pouvons pas savoir si l'auteur parle de la question didactique sur *l'attitude des étudiants* ou d'un document qu'il traite. Il n'y a pas de spécialisation dans ces deux cas, et la reformulation ne peut pas avoir lieu.

Dans l'exemple [3'], l'utilisation du marqueur *à savoir* ne fonctionne pas et empêche même la compréhension de l'énoncé.

- [3'] Enfin, pour le WWF (2001), l'écotourisme « simple » se mue en « écotourisme communautaire », **à savoir**, selon laquelle la communauté locale contrôle de façon significative son développement et sa gestion tout en y étant impliquée, et dans laquelle une proportion importante des bénéfices reste au sein des communautés. » [geo_09]

Dans cet exemple, *à savoir* ne peut pas fonctionner au plan sémantique. À notre avis, l'auteur cherche plutôt à expliquer la manière dont *l'énoncé simple se mue en écotourisme communautaire* qu'à le spécifier. L'apparition du marqueur *à savoir* nous paraît superflue dans cet exemple.

Ces quelques exemples d'utilisation des marqueurs discursifs dans les mémoires d'apprenants nous révèlent quelques erreurs fréquentes d'utilisation des MD chez les étudiants vietnamophones. Nous constatons que les étudiants rencontrent moins de problèmes de syntaxe que de difficultés sémantiques, ce qui peut s'expliquer par des propriétés syntaxiques régulières pour ces marqueurs. En revanche, il n'y a pas de régularité sémantique, et les étudiants n'ayant pas reçu d'enseignement sur ces aspects, maîtrisent difficilement ces points. Ces résultats permettent de confirmer que l'enseignement devrait s'orienter davantage sur les aspects sémantiques, car les étudiants maîtrisent en général les aspects syntaxiques.

Nous avons abordé jusqu'ici uniquement les expressions polylexicales, mais nous n'excluons évidemment pas le rôle de l'enseignement/apprentissage des monolexicaux en classe de français. En effet, ces derniers sont mieux abordés dans les méthodes de FLE tandis que les séquences lexicalisées à fonction discursive, en particulier les marqueurs à valeur métadiscursive⁴⁸, le sont beaucoup moins. Dans l'apprentissage du français, les apprenants ont besoin à la fois de ces expressions polylexicales et des unités monolexicales. C'est pourquoi, afin de développer les compétences discursives et rédactionnelles chez les apprenants, les réflexions didactiques ainsi que les activités que nous proposerons dans la partie suivante doivent prendre en compte tous ces éléments. Comme nous préconisons dans ce travail l'utilisation du corpus d'articles scientifiques pour l'enseignement/apprentissage des MD, nous allons maintenant décrire le rôle que peut jouer le corpus dans l'enseignement/apprentissage du français en général, et des MD en particulier.

6.2. Linguistique de corpus et enseignement/apprentissage du français

L'utilisation du corpus dans l'enseignement/apprentissage des langues est devenue un domaine très en vogue depuis les années 1990, particulièrement depuis les premiers travaux de Tim Johns (1991) sur le DDL *Data-Driven Learning* ou « apprentissage guidé par les données » (traduction littérale). Le DDL prend sa source dans l'informatique et s'impose comme une nouvelle approche en didactique des langues, avec la particularité d'allier des applications pédagogiques de la linguistique de corpus entre elles. Selon Johns (1991), le

⁴⁸ Cf. Chapitre 4 Typologie des séquences lexicalisées à fonction discursive.

DDL rend les apprenants plus autonomes dans leur apprentissage et leur permet de développer des stratégies de découverte par le fait qu'ils peuvent apprendre à apprendre (« learn to learn ») (*ibid.* : 1).

Nous retenons en français la notion proposée par Boulton (2011) : l'ASC, l'*apprentissage sur corpus*. Cette notion peut sembler un peu large par rapport à l'idée originale du DDL, mais elle sert à éviter la nuance entre *data-driven learning* et *data-based learning* (Boulton, 2008 : 5). Alors que l'approche par *corpus-based learning* utilise le corpus pour tester les idées existantes, l'approche par *corpus-driven approach* tente d'écarter toutes les intuitions pour donner un nouveau regard à l'appui des données (Tognini-Bonelli, 2001). Selon Boulton & Tyne (2014), la notion de l'ASC renvoie à « toute utilisation de données tirées de façon explicite d'un corpus pour exploitation par des non natifs à des fins d'apprentissage ou d'aide à l'expression écrite, la traduction, etc. » (*ibid.* : 145). Pourtant, à notre avis, l'ASC peut être utile même pour les locuteurs natifs. Même si tout ne peut pas être fait à partir du corpus, c'est à l'enseignant d'apporter toutes les explications de certaines notions, le corpus n'étant qu'un moyen pour apprendre à chercher. Il peut favoriser chez les apprenants une meilleure prise de conscience linguistique (Cavalla & Loiseau, 2013), dans la mesure où cette approche leur permet de développer un apprentissage en autonomie et une réflexion inductive comme le confirment plusieurs auteurs (Boulton, 2008, 2010 ; Aston, 2002).

Comme nous envisageons d'orienter l'enseignant et les apprenants vers un travail sur les corpus, nous chercherons à comprendre dans un premier temps les apports et les problèmes liés à l'introduction d'un tel outil en classe de langue. Puis, dans un deuxième temps, nous tenterons de définir le rôle de l'enseignant et de détecter les stratégies d'enseignement qui peuvent être mises en place. Ces précisions nous donneront des données complémentaires et nous guideront dans nos réflexions didactiques.

6.2.1. Apprentissage sur corpus

Avant de poser la question concernant la mise en place et l'utilisation de corpus en classe de langue, nous nous intéresserons aux avantages que peut apporter le travail sur corpus pour l'enseignement/apprentissage des langues. Nous cherchons à répondre successivement aux deux questions suivantes : comment les apprenants réagissent-ils face à

l'introduction d'un nouvel outil dans la classe de langue ? Existe-t-il des difficultés à l'intégration de corpus dans le programme d'enseignement/apprentissage ?

En ce qui concerne l'attitude des apprenants vis-à-vis de l'utilisation du corpus en classe de langue, plusieurs études montrent l'intérêt des apprenants face à un tel outil (Landure, 2013 ; Boulton & Tyne, 2014). Landure (2013) effectue deux expérimentations afin de comparer l'influence de l'utilisation d'un dictionnaire bilingue anglais-français, d'un dictionnaire unilingue et du corpus British National Corpus (BNC) dans l'apprentissage du phénomène linguistique des noms indénombrables en anglais. Le public se compose de deux groupes d'étudiants en 1^{re} année en Techniques de Commercialisation dans un Institut Universitaire de Technologie à une année d'intervalle. Les expérimentations montrent une efficacité dans la prise de conscience du phénomène linguistique étudié grâce au travail sur corpus. Plus particulièrement, le groupe qui bénéficiait d'une formation à l'utilisation des ressources obtient des résultats linguistiques plus intéressants que l'autre groupe, d'où l'importance d'une sensibilisation et d'une formation préalable aux phénomènes linguistiques auprès des étudiants. En outre, le bilan de plusieurs études sur l'apprentissage sur corpus que dressent Boulton & Tyne (2014) est plutôt positif dans l'ensemble. En effet, ils constatent un « gain d'autonomisation, développement de comportements cognitifs, responsabilisation dans l'apprentissage, capacité de travailler avec les données authentiques » (*ibid.* : 161).

Malgré les résultats positifs et encourageants de la part des apprenants dans l'utilisation des corpus, des réticences sur ce que peut apporter cet outil pour l'apprentissage existent toujours. Cavalla & Loiseau (2013) remarquent que de nombreux enseignants de FLE ignorent encore l'utilité d'introduire un corpus en classe de langues. Nombreux sont ceux qui ignorent aussi ce que sont ces corpus. Cela est peut-être lié au fait que les outils intégrant les corpus ne proposent pas encore une exploitation didactique directe des données. Ils constatent également une absence de formations des enseignants à l'exploitation des corpus en classe de FLE, ce qui n'est pas le cas pour l'anglais langue étrangère.

Afin d'expliquer une telle absence d'utilisation des corpus dans la classe de langue, Gilquin & Granger (2010) tentent de relever différentes difficultés qui viennent des enseignants et aussi des apprenants. Une des raisons de la non-utilisation du corpus de la part des enseignants est liée au manque de connaissances sur les possibilités que le corpus peut offrir à l'enseignement/apprentissage. À l'instar de Mauranen (2004), nous soulignons tout particulièrement l'importance d'une formation des enseignants : « Avant que les apprenants acquièrent de bonnes compétences avec les corpus, les enseignants doivent se les approprier

eux-mêmes en premier lieu » (notre traduction) (Mauranen, 2004 : 100, cité par Guilquin & Granger, 2010 : 366).

Une autre raison relevée par Gilquin & Granger (*ibid.*), et qui ne semble pas anodine, est relative aux préoccupations des enseignants dans le changement de la nature de l'apprentissage, essentiellement des rôles des enseignants et apprenants face à un nouvel outil de travail. Il nous semble important de faire prendre conscience à l'enseignant du rôle qu'il joue dans l'élaboration du contenu d'enseignement. Johns (1991) insiste également sur le rôle crucial de l'enseignant pour guider les apprenants et préparer les activités. Ce sont eux qui font la sélection des exemples les plus adéquats en prenant en compte le niveau et le besoin des apprenants. L'enseignant est le premier acteur à décider de la réussite d'un processus d'enseignement/apprentissage. Des formations à l'utilisation des corpus pour l'enseignement s'avèrent donc indispensables.

Du côté des apprenants, Gilquin & Chambers (2010) mentionnent une autre difficulté, plutôt cognitive. Le fait de mettre l'apprenant devant l'ordinateur et de lui demander de relever les règles précises n'est pas chose facile. Sans les explications de l'enseignant, les apprenants ont parfois du mal à formuler les règles. C'est pourquoi en fonction de la progression des apprenants, l'enseignant peut intervenir afin de les aider à synthétiser le contenu d'apprentissage, mais aussi expliquer plus clairement les objectifs linguistiques et/ou communicatifs mis en jeu, et relever les traits syntaxiques et sémantiques d'un phénomène linguistique.

Une dernière difficulté compliquant l'utilisation des corpus chez les apprenants selon Gilquin & Chambers (*op.cit.*) concerne le contenu de l'ASC. Les apprenants risquent de se sentir intimidés face à un nombre d'occurrences très élevé à l'issue des requêtes sur corpus. Alors que les apprenants de niveau débutant et novice dans l'utilisation des corpus se sentent vite « ennuyés/embarassés » s'ils doivent travailler sur une cinquantaine d'occurrences ou plus, les étudiants de niveau avancé ne voient pas forcément l'intérêt de travailler sur un nombre d'occurrences aussi important. Cela risque d'augmenter la surcharge cognitive des apprenants. Tout cela souligne d'autant plus le rôle de l'enseignant dans la sélection préalable des exemples extraits du corpus susceptible d'intéresser les apprenants en s'appuyant sur leurs objectifs de travail et leur niveau de langue.

Par ailleurs, le recours aux données authentiques correspond aux recommandations méthodologiques du CECRL en permettant aux étudiants de s'exposer aux textes authentiques

en L2. Cependant, l'introduction du corpus en classe de langue rencontre quelques difficultés liées à la masse d'informations ou au scepticisme face à une modification de la nature de l'enseignement/apprentissage. Les réflexions didactiques que nous allons proposer vont tenter de prendre en compte ces problèmes afin d'aider les apprenants à tirer profit des avantages que peut apporter l'apprentissage sur corpus.

L'introduction des corpus dans le programme d'enseignement/apprentissage demande une mise en œuvre de nouvelles stratégies d'enseignement. Analysons maintenant quelques stratégies qu'il est nécessaire de développer pour l'intégration de ce nouvel outil d'enseignement/apprentissage en classe de langue.

6.2.2. Stratégies d'enseignement et rôle des enseignants

Dans ce travail, nous souhaitons sensibiliser les enseignants et les apprenants à l'utilité et à la plus-value de l'utilisation d'une base de données textuelle dans l'apprentissage du français, plus précisément dans l'apprentissage des MD dans les écrits scientifiques. Fligelstolle (1993) regroupe en trois catégories les principaux usages du corpus à savoir : 1/ *teaching about* (enseigner la linguistique et la méthodologie) ; 2/ *teaching to exploit* (enseigner aux étudiants à exploiter le corpus) ; 3/ *exploiting to teach* (exploiter le corpus pour enseigner) (*ibid.* : 98). Si le premier usage renvoie plutôt aux connaissances de base du domaine de la linguistique de corpus, sur les théories, les objectifs, etc., les deux autres usages évoquent selon nous les différentes activités mises en place pour développer les compétences linguistiques et métalinguistiques chez les enseignants et les apprenants. D'une part, pour le deuxième usage, les étudiants apprennent différentes techniques, les requêtes pour travailler sur un corpus. Et, d'autre part, le troisième usage est plutôt orienté vers l'enseignant, notamment sur les meilleures façons d'exploiter le corpus pour des objectifs d'enseignement. Si nous considérons que l'utilisation sur corpus se déroule en deux temps, il s'agit dans un premier temps d'une préparation par l'enseignant en amont du travail en classe. Celui-ci joue le rôle de médiateur qui va programmer les parcours de travail, préparer des activités et sélectionner des exemples qui correspondent au niveau et aux besoins des apprenants. Dans un deuxième temps, les apprenants travaillent sur le corpus sous le guidage de l'enseignant (Cavalla, 2014).

L'intégration du corpus en classe de langue remet en question le rôle de l'enseignant et les stratégies d'enseignement à mettre en place. Dans le CECRL, « est considéré comme

stratégie tout agencement organisé, finalisé et réglé d'opérations choisies par un individu pour accomplir une tâche qu'il se donne ou qui se présente à lui » (CECRL, 2001 : 15). Cette définition du CECRL renvoie selon nous plutôt aux stratégies d'apprentissage. Dans ce travail, nous nous intéressons tout particulièrement aux stratégies d'enseignement, autrement dit, à ce que l'enseignant peut mettre en place pour optimiser l'enseignement. En effet, l'enseignant doit adapter sa pratique d'enseignement à l'introduction d'un nouvel outil de travail. Nous nous penchons plus précisément sur le troisième usage du corpus, *exploiting to teach* proposé par Fligelstolle (1993). Plusieurs études sur l'ASC insistent également sur l'importance de la formation des formateurs (Delahaie & Flament-Boistrancourt, 2013), car les enseignants jouent un rôle crucial dans la mise en place de l'ASC. C'est pourquoi nous insistons sur la formation des enseignants pour l'utilisation du corpus en tant qu'outil d'apprentissage.

Les enseignants ont besoin d'être guidés dans l'élaboration d'une séquence d'apprentissage des MD qui s'appuie sur un corpus, y compris pour la sélection des données et l'organisation du programme (Boulton & Tyne, 2014). En effet, pour les documents authentiques introduits en classe qui ont des formats connus (image, enregistrements oraux, vidéo, etc.), l'enseignant sait comment les analyser pour les enseigner, car ceci est connu de longue date. En revanche, pour un corpus, il faut former l'enseignant à l'utilisation de ce nouveau « document authentique » pour qu'il puisse savoir comment extraire les données et ensuite savoir les lire en vue d'une exploitation didactique, ce que les enseignants de FLE ne savent peut-être pas encore faire, car ils n'ont, pour le moment, pas reçu de formation dans ce domaine. Boulton & Tyne (2014) confirment l'intérêt d'un apprentissage à partir des exemples préétablis par l'enseignant. Selon eux, « le travail direct sur le corpus électronique est parfois perçu comme étant moins intéressant ou efficace que le travail avec des données présélectionnées par l'enseignant et appuyées par des activités ciblées » (*ibid.* : 169).

Outre cela, il faut aussi guider l'enseignant dans la sélection des exemples. L'enseignant doit être sensibilisé aux spécificités des écrits scientifiques, aux fonctions d'écriture, aux structures discursives, sur le fonctionnement des MD notamment. C'est pourquoi une formation de l'enseignement des MD peut être envisageable pour expliquer aux enseignants les différentes fonctions de ces éléments dans les écrits. Il est nécessaire, à notre avis, de leur transmettre les différentes approches existantes pour optimiser l'enseignement des MD en particulier, et d'autres éléments du lexique en général.

En ce qui concerne l'organisation de l'apprentissage, l'enseignant doit prévoir la manière dont se déroulent les exercices et anticiper les éventuelles questions des apprenants (Boulton & Tyne, 2014). L'introduction du corpus dans l'enseignement/apprentissage du français ne modifie pas l'idée d'un apprentissage centré sur les apprenants, car l'enseignant reste leur guide. Dans notre typologie des séquences lexicalisées à fonction discursive (Cf. Chapitre 4), nous ne recensons que les MD les plus fréquents des articles scientifiques. Cette liste n'est sans doute pas exhaustive, mais elle peut être considérée comme un guide qui apporte à l'enseignant des éléments essentiels pour élaborer ses parcours pédagogiques. Certains MD n'apparaissent pas dans cette liste en raison de leur faible fréquence dans notre corpus. Ils ne doivent pas pour autant être exclus de l'apprentissage. Pour un travail sur un texte long, l'enseignant devra indiquer aux apprenants ces éléments. De plus, même si nous traitons uniquement des expressions polylexicales au plan linguistique, nous n'écartons pas non plus les expressions monolexicales dans les activités didactiques. Prenons l'exemple de la reformulation : l'enseignant peut repérer dans un article scientifique des marqueurs *en bref*, *pour conclure*, *pour finir*, etc. qui servent à introduire une reformulation, mais ces marqueurs n'apparaissent pas dans notre typologie. Il est donc nécessaire que l'enseignant précise également ces unités lexicales aux apprenants en les faisant travailler sur un texte précis. Les fiches pédagogiques que nous allons présenter à la fin de ce chapitre (Cf. 6.3.2 Propositions didactiques, page 241) peuvent être considérées comme un exemple d'intégration d'extraits de corpus ou des concordanciers dans un cours de langue.

Un autre point est à aborder avec l'enseignant : il faut le sensibiliser à l'intérêt du corpus en classe de langue. Le corpus permet à l'enseignant d'expliquer les différents emplois des éléments lexicaux en les observant, ce qui est indispensable pour l'enseignant natif, en encore plus pour l'enseignant non natif. Ce dernier ne peut que difficilement recourir à son intuition, et l'utilisation du corpus lui donne donc un certain confort. Par exemple, concernant la question de la synonymie, l'utilisation du corpus sert à relever les propriétés sémantiques qui permettent de distinguer un MD d'un autre.

Ainsi, nous insistons également sur la mise en place des formations vis-à-vis des enseignants pour développer les compétences d'utilisation du corpus et également pour traiter de différentes approches d'enseignement des MD. Ces formations aux enseignants ont pour objectif de les aider à prendre en compte avant tout l'intérêt ainsi que l'efficacité d'un apprentissage sur corpus pour pouvoir le transmettre aux étudiants. Pour que les apprenants apprécient l'apprentissage sur corpus, il faudrait avant tout que les enseignants prennent

plaisir à la découverte qu'apporte le travail sur corpus, et qu'ils prennent conscience de l'importance de l'enseignement de certains phénomènes linguistiques indispensables à certains genres comme l'écrit scientifique.

Il nous semble important à tout niveau, et particulièrement au niveau universitaire, d'aider les étudiants à acquérir une méthode de travail de façon à ce qu'ils puissent la réinvestir pour l'apprentissage d'autres contenus universitaires. Les réflexions didactiques ainsi que les activités didactiques que nous allons aborder dans la partie suivante s'adressent essentiellement aux enseignants. Elles tentent d'apporter d'une part quelques clarifications pour l'enseignement/apprentissage des MD, et, d'autre part, pour celui d'autres éléments linguistiques.

6.3. Réflexions didactiques : vers des propositions

Nos réflexions didactiques dans cette partie s'intègrent dans le cadre du Français sur Objectif Universitaire (FOU) défini par Mangiante & Parpette (2012) comme « une démarche englobante déclinant celle du FOS qui suppose la mise en place d'un véritable dispositif impliquant l'Université et ses partenaires institutionnels » (*ibid.* : 163). L'enseignement d'éléments phraséologiques dans les écrits scientifiques s'inscrit dans cette démarche dès l'instant où il est nécessaire d'inclure des cours d'aide à l'écriture scientifique pour des étudiants non natifs du français. Les exigences de l'écriture universitaire française font appel à des compétences discursives dans les écrits universitaires attendus (mémoires, rapports, etc.).

Il nous semble nécessaire d'envisager trois objectifs didactiques : 1/ sensibiliser les enseignants comme les apprenants à l'importance de la prise en compte des MD comme unités d'apprentissage ; 2/ montrer comment les MD participent à la structuration du texte ; 3/ faire acquérir auprès des étudiants des spécificités d'un genre universitaire que sont les articles scientifiques. Pour cela, nous souhaitons proposer un prototype d'enseignement/apprentissage des marqueurs discursifs adaptable aux étudiants en SHS, dans l'intention de les aider à développer leurs compétences rédactionnelles et discursives. Le corpus devient à la fois le terrain de travail pour effectuer des analyses linguistiques, mais également une source de données dont l'enseignant peut s'inspirer pour mettre en place différentes activités didactiques.

6.3.1. Réflexions didactiques

À ce jour, les réflexions approfondies sur l'enseignement des MD auprès des étudiants ne sont pas encore suffisantes et les enseignants ont eux aussi besoin d'être guidés dans l'élaboration d'un parcours pédagogique. C'est pourquoi nous souhaitons mettre en avant quelques principes importants à prendre en considération pour l'enseignement de ces éléments linguistiques grâce au corpus. Il peut s'agir de différentes stratégies d'enseignement et également des ressources que nous mettons à leur disposition.

Comme les analyses syntactico-sémantiques des MD effectuées dans le chapitre 5 ne se prêtent pas à une exploitation directe par les enseignants, nous pouvons considérer les lignes qui suivent comme des suggestions pour la transposition didactique de ces analyses linguistiques. Cette partie s'organise de la manière suivante : dans un premier temps, nous développerons différentes approches que l'enseignant peut mettre en place pour enseigner les MD et accéder aux écrits scientifiques. Dans un deuxième temps, nous proposerons quelques exemples d'activités en tenant compte de ces approches. Nous pensons qu'il serait intéressant d'introduire les différentes approches didactiques en classe de langue. En raison de la complexité sémantique de notre objet de travail, nous supposons qu'il ne faut pas envisager une seule approche en début d'apprentissage.

Outre l'objectif linguistique, nous souhaitons également proposer une méthodologie d'aide à la rédaction scientifique. Pour atteindre nos objectifs, nous proposons d'associer les deux entrées suivantes : 1/ une entrée conceptuelle, avec la combinaison des approches inductive et onomasiologique, nous semble très propice à l'enseignement/apprentissage des phénomènes linguistiques à l'aide du corpus ; 2/ une entrée lexicale associée à des approches déductive et sémasiologique s'avère très efficace pour la consolidation des connaissances. Nous intégrons également des activités sur les aspects discursifs qui permettront à l'apprenant de s'approprier les spécificités des écrits scientifiques, et des fonctions rhétoriques.

6.3.1.1. *Approche inductive et approche onomasiologique*

a. Approche inductive

L'approche inductive décrit un type de raisonnement qui va du particulier au général, des « données particulières (faits, expériences, énoncés) à des propositions plus générales, de cas particuliers à la loi qui les régit, des effets à la cause, des conséquences au principe, de l'expérience à la théorie » (d'après le TLFi). L'approche inductive est très utilisée pour enseigner les phénomènes grammaticaux ; le CECRL préconise d'ailleurs cette démarche pour développer la compétence grammaticale des apprenants. Ces derniers travaillent sur des « données grammaticales » issues de documents authentiques pour ensuite relever leur fonction et leur signification. Cette approche commence à susciter l'intérêt des chercheurs dans de nombreux domaines sur l'exploitation de corpus au service de l'enseignement/apprentissage (Chambers, 2010 ; Boulton & Tyne, 2014) et sur l'acquisition langagière (Colletta, 2004). Colletta (2004) privilégie l'entrée inductive qui est « davantage propice à des prises de conscience de ce type [prendre conscience du fonctionnement de sa langue et pouvoir verbaliser les phénomènes linguistiques] que l'entrée déductive où la règle est simplement formulée par l'enseignant et illustrée à l'aide d'un exemple choisi pour son caractère typique » (*ibid.* : 311).

En fait, l'approche inductive prend source dans le constructivisme, essentiellement dans les travaux de Vygotsky (1896-1934) pour qui l'apprentissage pilote le développement. L'enseignement constructiviste met l'accent sur les activités que l'apprenant peut accomplir seul pour construire de nouvelles règles en s'appuyant sur son expérience, ses connaissances préalablement acquises (Lasnier, 2000). L'approche inductive est également un des principes sous-jacents du DDL que propose Johns (1991) quand il suggère de considérer les apprenants comme des « détecteurs ». Ils deviennent plus responsables et autonomes dans leur apprentissage.

En outre, l'approche inductive permet aux apprenants d'endosser un nouveau rôle et de se trouver au cœur de leur apprentissage (Boch & Buson, 2012 ; Blondel, 2012). Ils sont donc plus motivés comme le constatent Boch & Buson (*ibid.*), quand ils jouent le rôle de chercheur ou linguiste dans leur recherche (essentiellement de niveau avancé) et sont sollicités pour découvrir le fonctionnement de chaque phénomène linguistique. La démarche inductive a l'avantage de placer « les étudiants dans une posture réflexive, rompant ainsi avec une attitude passive consistant à recevoir et à (plus ou moins bien) utiliser une forme trop peu questionnée » (*ibid.* : 47). En somme, une approche inductive correspond bien à notre objectif puisqu'elle favorise une meilleure prise de conscience linguistique chez les apprenants. L'observation et le repérage des MD dans des contextes authentiques facilitent une bonne

mémorisation chez les apprenants car ceux-ci peuvent relever eux-mêmes le fonctionnement syntaxique et sémantique des MD. Il est à rappeler que nous insistons tout particulièrement sur le rôle de l'enseignant dans l'élaboration du programme de travail et dans l'organisation du cours. Toute la partie explicative de certaines notions doit passer par l'enseignant, l'apprentissage évoluant donc dans une approche inductive guidée.

Nous préconisons un enseignement/apprentissage des MD par l'approche inductive dès le niveau B1 pour ce type d'écrit spécifique. Cette approche permet aux apprenants de relever des régularités syntaxiques et sémantiques des MD et de formuler des règles, ce qui facilite une bonne mémorisation par rapport à une méthode habituelle où l'enseignant donne des règles et fait travailler les apprenants sur quelques exercices. En ce qui concerne les exercices, l'enseignant peut utiliser des exercices d'observation et de repérage ou encore des exercices lacunaires pour la mise en comparaison des MD. Une telle approche se marie bien, nous semble-t-il, avec l'approche onomasiologique reprise de la linguistique en didactique, car dans les deux cas, il est attendu une sorte de conceptualisation des connaissances par l'apprenant. Voyons ci-dessous comment allier les deux entrées.

b. Approche onomasiologique

Dans notre travail, nous préconisons l'utilisation de l'approche onomasiologique dans un premier temps pour l'enseignement/apprentissage des MD. Comme l'analyse du corpus d'apprenants montre que les erreurs des apprenants se trouvent essentiellement sur le plan sémantique, c'est donc sur cet aspect que nous souhaitons attirer leur attention. Nous souhaitons sensibiliser aussi bien les enseignants que les apprenants sur les différentes fonctions discursives spécifiques des écrits scientifiques (reformulation, introduction du contenu thématique, etc.). Cette approche onomasiologique nous semble très adaptée pour développer les compétences rédactionnelles, d'autant plus qu'elle permet aux apprenants d'accéder aux différentes fonctions d'écriture dont ils ont besoin dans la rédaction.

Par l'enseignement/apprentissage des MD, nous souhaitons aider les étudiants à apprendre à structurer leur texte et à mieux organiser les idées. Avec l'approche onomasiologique, au lieu de donner aux apprenants une liste de mots, l'enseignant choisit la fonction rhétorique qu'il souhaite aborder selon leur niveau, leurs besoins et leurs objectifs d'apprentissage et ensuite, apparaît la question de savoir quels mots utiliser pour énoncer cette fonction rhétorique. Cette approche peut aussi être mise en œuvre par des activités de

compréhension écrite et de production écrite. Pour la compréhension écrite, l'enseignant peut commencer par des exercices de repérage : repérer dans le texte les marqueurs servant à introduire une thématique ou à reformuler, etc. L'enseignant peut ensuite recourir à l'approche inductive en demandant aux apprenants de relever eux-mêmes les valeurs d'usage de chaque MD, en mettant à leur disposition de plus amples exemples. Pour la production écrite, l'enseignant peut proposer des activités qui portent essentiellement sur une fonction discursive, par exemple, utiliser les marqueurs servant à introduire une succession d'idées pour développer un argument, rédiger une introduction en précisant le plan de développement du texte, etc. Il est à noter que nous allons mettre en œuvre ces considérations méthodologiques dans les fiches pédagogiques présentées dans la partie **6.3.2**.

Rappelons que notre typologie des séquences lexicalisées à fonction discursive a été établie selon une approche onomasiologique (Cf. *Annexe 1*). Pour faciliter l'apprentissage des MD, l'enseignant peut commencer par ce qui est le plus transparent au niveau sémantique comme les marqueurs métadiscursifs (les marqueurs servant à introduire un contenu, les marqueurs d'exemplification, les marqueurs de topicalisation, etc.), pour aborder plus tard les marqueurs argumentatifs. Les raisons sont plutôt d'ordre cognitif. Dans la catégorie des marqueurs métadiscursifs, les apprenants peuvent relever d'autres éléments monolexicaux dans l'écrit scientifique (*enfin, ensuite, premièrement, deuxièmement*, etc.) qui n'apparaissent pas dans notre liste et qui apparaissent plus souvent dans les méthodes de FLE. Ces éléments, dont le sens est facilement décodable, sont donc souvent connus des apprenants. L'enseignant peut partir de ces éléments connus pour passer à d'autres expressions polylexicales inconnues et parfois complexes. En effet, si le cadre sémantique est posé via des éléments connus, l'entrée avec des MD inconnus devrait en être facilitée. Ajoutons que les MD argumentatifs (explication, concession, opposition) sont complexes tant au plan sémantique que syntaxique. Nous suggérons alors de les aborder dans un deuxième temps, une fois que les apprenants auront acquis des compétences de travail.

Prenons maintenant l'exemple de l'introduction des approches inductive et onomasiologique pour l'enseignement/apprentissage des marqueurs de reformulation. En travaillant sur les marqueurs de reformulation, l'enseignant peut sensibiliser les apprenants à leur fonction pragmatique, c'est-à-dire que la reformulation manifeste le souci d'éclairer le discours pour le lecteur. L'enseignant peut commencer par les marqueurs les plus fréquents et les plus productifs comme les marqueurs servant à clarifier, aborder ensuite les MD servant à récapituler et enfin, les MD de reconsidération. La raison est plutôt d'ordre pédagogique :

d'un côté, les marqueurs servant à clarifier (*c'est-à-dire*, *à savoir*, *autrement dit*, *en d'autres termes*) sont les plus utilisés par les locuteurs natifs ; d'un autre côté, leur portée est relativement courte par rapport à celle des MD d'autres sous-groupes. Ceci semble important au plan didactique, car pour une découverte de tels phénomènes linguistiques, autant ne pas surcharger les apprenants de phrases complexes. Plus la portée est courte dans un premier temps, plus cela incite les apprenants à les utiliser. Si dès le départ on présente quelque chose de trop complexe, les découragements seront fréquents. Il est donc plus judicieux de faire travailler les apprenants sur des contextes assez courts pour les aider à acquérir quelques compétences de travail sur corpus et, également, à s'habituer au lexique des écrits scientifiques. Les autres MD, surtout ceux de reconsidération qui servent à introduire le point de vue de l'auteur, sont plus complexes d'utilisation dans un premier temps, nous semble-t-il. Pour illustrer cela, nous avons créé un exercice pour l'introduction des marqueurs de reconsidération (Cf. **Fiche pédagogique 2**, page 258)

Concernant les aspects syntaxiques des MD, pour les marqueurs de récapitulation (*en somme*, *en résumé*), le contexte gauche du MD peut être relativement long. L'enseignant peut aider les apprenants à relever différentes idées ou arguments dans le contexte gauche du MD pour détecter par la suite le rapport établi entre deux contextes par le MD.

Comme nous l'avons évoqué brièvement dans les analyses linguistiques des MD de reformulation, l'enseignement des MD via la synonymie n'est pas forcément efficace, par exemple *à savoir* ou *autrement dit* ne peuvent pas être remplacés d'une manière automatique par *c'est-à-dire*. Ces marqueurs qui se trouvent dans la même catégorie sémantique et ne partagent pas tout à fait les mêmes propriétés sémantiques sont appelés « presque synonymes » (Kamber, 2011) ou « parasynonymes » (Picoche, 1993). Dans l'enseignement du vocabulaire, l'enseignant doit tenir compte de ce phénomène.

D'après notre étude de corpus, chaque MD se caractérise par des valeurs d'usage spécifique, et certains MD peuvent se substituer uniquement dans quelques contextes précis. Afin d'aider les apprenants à acquérir les propriétés spécifiques de chaque MD, nous suggérons de mettre en comparaison les MD sémantiquement proches. Plus précisément, nous proposons d'enseigner les MD par petits groupes de « parasynonymes » pour mieux faire ressortir le fonctionnement de chaque MD. Dans chaque groupe, l'enseignant peut faire travailler les apprenants de prime abord sur chaque MD individuel puis par exemple : 1/ mettre en comparaison les MD d'un même sous-groupe (comme *c'est-à-dire*, *à savoir*, *autrement dit*, *en d'autres termes*) ; 2/ regrouper les MD qui sont souvent définis comme

synonymes de l'un de l'autre dans les dictionnaires : *en somme/somme toute ; en somme/au total ; en définitive/en fin de compte/au fond*.

Se trouvant toujours dans une approche inductive, les apprenants vont travailler sur un ensemble d'exemples de différents marqueurs et devront relever les valeurs sémantiques de chaque MD. Pour approfondir, l'enseignant propose des exercices de substitution afin de mieux voir dans quel cas tel MD peut être remplacé par tel autre. Ces activités permettent selon nous d'éviter le recours automatique à la notion de *synonymie* dans l'enseignement du lexique. L'enseignant peut se servir des tableaux récapitulatifs des caractéristiques de chaque marqueur, ainsi que des tableaux de comparaison que nous avons mis en place dans le chapitre 5. Ces tableaux donnent les propriétés syntaxiques et sémantiques de chaque MD, ainsi que ses spécificités, pour les aider dans la préparation des cours. (un exemple sur l'introduction des marqueurs souvent considérés comme synonymes *c'est-à-dire* et *à savoir* se trouve dans la partie 6.3.2).

Pour une approche onomasiologique, l'enseignant peut également se servir des tableaux récapitulatifs des propriétés sémantiques que nous avons mis en place dans le chapitre 5 pour proposer une entrée par valeur sémantique. Par exemple, il peut faire travailler les apprenants sur les valeurs sémantiques des MD : marqueurs servant à exprimer une valeur résumante (*c'est-à-dire*) ou une valeur énumérative (*à savoir*), etc. Pour les activités de production écrite, l'enseignant peut envisager des activités de reformulation et demander aux apprenants de reformuler les idées en utilisant les MD étudiés.

En outre, l'enseignant peut envisager de faire travailler les apprenants sur les concordanciers des MD ou des extractions du corpus. Ceux-ci sont susceptibles de les aider à prendre conscience du fonctionnement du MD pour pouvoir le réutiliser par la suite. L'utilisation du concordancier dans l'apprentissage des langues prend source au sein du DDL⁴⁹ évoqué précédemment. Le concordancier peut apporter de nombreux avantages pour l'enseignement/apprentissage, notamment permettre à l'apprenant de travailler sur des exemples authentiques, et donc de s'approcher du lexique dans les écrits scientifiques. Nous allons présenter dans la partie suivante quelques exemples d'utilisation des concordanciers pour enseigner les MD. Ces concordanciers sont extraits de notre corpus Scientext (Cf.6.3.2 Exercices avec des concordanciers). De plus, nous mettons à disposition de l'enseignant et

⁴⁹ Cf.6.2.1 pour la description du DDL, page 14.

des apprenants un dictionnaire basé sur le corpus Dicorpus que nous allons présenter plus en détail ultérieurement.

Nous proposons de mettre en œuvre l'approche inductive et onomasiologique par un aspect spécifique des écrits scientifiques, à savoir leurs fonctions rhétoriques.

c. Aspect rhétorique

Les considérations méthodologiques que nous venons d'aborder concernent uniquement l'enseignement/apprentissage des MD. Comme ce travail s'intègre dans la démarche du FOU et que notre objectif final est d'aider les apprenants à mieux organiser leur écrit universitaire, les MD font partie, selon nous, des moyens lexicaux susceptibles de les sensibiliser à la structuration du texte. L'enseignement/apprentissage des MD doit également tenir compte des « spécificités architecturales et linguistiques » (Cavalla, 2006) dans les écrits scientifiques. L'approche par l'aspect rhétorique, (à la fois onomasiologique et inductif), représente pour nous une autre manière d'aborder les MD et d'accéder à ce genre d'écrits.

Dans cette recherche, les apprenants vont travailler sur un corpus d'articles scientifiques qui constituent un genre nouveau auquel les apprenants de Licence et de Master ne sont pas encore accoutumés. Rédiger un écrit scientifique ne ressemble pas à la rédaction d'une dissertation ou d'un récit. Il existe des codes propres à ce genre que les apprenants doivent s'approprier. Ainsi, nous rejoignons Boulton & Tyne (2014) qui soulignent l'importance d'un travail d'analyse et d'observation des exemples authentiques pour ce genre d'apprentissage. En effet, en faisant travailler les apprenants sur des données extraites du corpus d'articles scientifiques, ils peuvent non seulement travailler les aspects linguistiques (le lexique, les MD), mais aussi les fonctions spécifiques de ce genre d'écrits. Ils peuvent s'initier au style d'écriture, à la manière dont l'auteur se positionne dans son écrit, et aussi à l'agencement des idées dans les différentes parties textuelles (introduction, développement, conclusion). Par exemple, dans l'introduction, l'auteur peut présenter le contexte et le thème de la recherche, les choix méthodologiques, la problématique, le plan, etc. Ce genre d'activité de repérage des différentes structures discursives s'avère indispensable pour développer les stratégies de lecture que nous allons aborder plus loin.

Une autre caractéristique importante des écrits scientifiques qu'il est nécessaire de présenter aux étudiants est leur fonctionnement « dialogique » (Grossmann & Tutin, 2010). Le lecteur d'un écrit scientifique se trouve dans un rapport entre pairs : « Pour écrire son

texte, l'auteur se forge une représentation relativement précise de ses lecteurs, qui sont aussi ses pairs, et c'est à partir de cet auditoire idéal qu'il se pose lui-même comme énonciateur » (*ibid.* : 2). La prise en compte des lecteurs peut se manifester plus ou moins implicitement, soit par des traces référant au lecteur parmi lesquelles l'usage des pronoms *nous/on*, soit par l'emploi d'un impératif de première personne du pluriel (*comparons à nouveau nos résultats*,⁵⁰ etc.). L'auteur peut également faire part de « la volonté de faire comprendre son intention au lecteur par un ensemble de marques de guidage », comme les marques métatextuelles ou les marques de reformulation (*ibid.* : 3).

En résumé, pour une écriture universitaire scientifique, l'enseignement/apprentissage ne s'arrête pas uniquement au niveau lexical, c'est-à-dire à l'acquisition de la fonction syntaxique et sémantique d'un élément linguistique. Il est important d'aider les apprenants à se familiariser avec différentes fonctions plus larges, notamment les fonctions rhétoriques spécifiques de ce genre d'écrit. Cet enseignement permettra aux apprenants de prendre en considération leurs lecteurs quand ils rédigent, ce que la plupart d'entre eux ignorent encore. De son côté, l'enseignant a le rôle de préciser cet aspect aux étudiants en l'introduisant en classe de langue. Il peut l'insérer au sein de l'approche onomasiologique puis sémasiologique comme présentée ci-après.

Comme nous l'avons vu dans le chapitre 5, les marqueurs de reformulation constituent un des moyens de guidage de l'auteur pour faciliter la compréhension du lecteur. D'autres marqueurs assument également cette fonction comme les marqueurs d'exemplification, de topicalisation, etc. (Cf. Annexe 1). Outre les MD, les « routines sémantico-rhétoriques » spécifiques aux écrits scientifiques (Tutin, 2013) sont un autre élément phraséologique qu'il est important de présenter aux apprenants. Elles sont « des énoncés stéréotypés renvoyant à des fonctions rhétoriques [...] comme la filiation scientifique indiquée par un auteur, la démarcation par rapport à un pair, l'évaluation des activités scientifiques des pairs, l'élaboration de preuves factuelles, par exemple *comme on peut le voir ci-dessous, on l'a vu ci-dessus*, etc. » (*ibid.* : 34). L'introduction des routines sémantico-rhétoriques est très importante pour la compréhension écrite et la production écrite. Pour la compréhension écrite, l'apprenant peut mieux repérer les fonctions rhétoriques liées à chaque routine sémantico-rhétorique. Pour la production écrite, l'utilisation de ces « segments déjà vus » peut aider à améliorer les techniques d'écriture (Mangenot, 1997) dès l'instant où ils permettent d'engager

⁵⁰ Exemple dans Grossmann & Tutin (2010).

l'étudiant en apprentissage dans un rapport entre pairs. Nous rassemblons donc ici les approches inductives, onomasiologiques et rhétoriques pour une aide à la lecture et à la production écrite scientifique telle que le préconise Lundquist (*op.cit.*).

Dans ce travail, nous souhaitons développer chez les apprenants les stratégies de lecture pour développer par la suite les stratégies de production. Nous envisageons de mettre en rapport une « lecture active » (Lundquist, 2013) et une « écriture dynamique » (« dynamics of composing ») (Flower & Hayes, 1980). En effet, les stratégies de production font écho aux stratégies de lecture. D'une part, les connaissances sur les spécificités architecturales et sur le genre que les apprenants auront acquises tout au long de leur apprentissage vont leur permettre de mieux organiser le texte et de mettre l'accent sur les idées essentielles. Ils peuvent s'inspirer des modèles de structuration du texte pour élaborer un plan de leur écrit. D'autre part, les apprenants peuvent mobiliser des moyens linguistiques (MD, collocation, « routines sémantico-rhétoriques », etc.) ainsi que des structures discursives pour les réinvestir dans leur écrit. En fait, l'analyse du corpus et les activités de compréhension écrite apportent des ressources lexicales et des savoirs langagiers indispensables à la production écrite. En outre, l'apprentissage des MD permet aux apprenants de comprendre comment les MD participent à l'organisation du texte pour mieux repérer les idées essentielles.

Pour une compréhension écrite « active », il est nécessaire de donner aux apprenants/étudiants des méthodes de travail, surtout quand ils utilisent un corpus composé d'articles venant de disciplines très variées. Certes, la question de la terminologie existe toujours, mais ces articles sont régis par les mêmes principes du genre « scientifique » comme nous l'avons vu dans le chapitre 3. Pour une compréhension écrite « active » (Lundquist, 2013), l'enseignant peut, en premier lieu, habituer les apprenants aux régularités et aux spécificités des articles scientifiques. Selon Lundquist (*ibid.*), pour développer la stratégie de lecture, il faut faire « prendre conscience aux lecteurs novices de textes académiques de ces principes généraux de lecture, afin que d'apprenants non experts dans une discipline choisie, ils acquièrent des stratégies de lecture leur facilitant l'accès au savoir nouveau des textes académiques, et les rendant finalement autant experts en lecture que dans leur domaine » (*ibid.* : VI). Nous sommes du même avis que Lundquist sur le fait qu'il est important d'enseigner aux apprenants une méthodologie de travail afin qu'ils deviennent autonomes dans leur apprentissage. Quand les apprenants ont davantage l'occasion de s'exposer à la langue (au niveau lexical et architectural), ils peuvent mettre en place les caractéristiques spécifiques de la langue en question. Pourtant, nous émettons une réserve quant au choix de

textes entiers à faire travailler aux apprenants. L'enseignant peut choisir des extraits venant de différentes disciplines pour leur faire observer les phénomènes linguistiques. Cependant, le choix d'un texte entier d'une discipline différente du domaine de compétence de l'apprenant est une autre chose. Nous allons encore revenir sur cette question ultérieurement quand nous explicitons le choix des textes dans nos fiches pédagogiques. Mais reprenons les stratégies proposées par Lundquist afin de comprendre ensuite comment les développer dans notre cas.

➤ **Stratégies de lecture *top-down***

Nous nous inspirons de deux processus proposés par Lundquist (2013) : le *top-down* (du haut vers le bas) et le *bottom-up* (du bas vers le haut) pour développer les compétences discursives des apprenants. Ces deux stratégies d'enseignement s'appuient d'un côté sur les connaissances des apprenants sur la langue et « sur la structuration et le déroulement de texte en général » (*ibid.* : V) (pour la stratégie de lecture *top-down*) et d'un autre côté, sur les unités lexicales grammaticales (la stratégie de lecture *bottom-up*). La stratégie *top-down* ressemble plutôt aux principes des approches inductive et onomasiologique, tandis que la stratégie *bottom-up* peut rejoindre les concepts des approches déductive et sémasiologique que nous verrons plus loin. Autrement dit, les approches inductive et onomasiologique favoriseront la mise en place de la stratégie *top-down*, et les approches déductive et sémasiologique permettront de développer la stratégie *bottom-up*. Nous aborderons la notion de stratégie *bottom-up* dans la partie suivante sur les approches déductive et sémasiologique et nous attarderons ici à décrire la stratégie de lecture *top-down*.

Les connaissances sur les structures discursives dans l'écrit scientifique permettent aux apprenants de développer les stratégies *top-down*. Prenons l'exemple des activités de compréhension et de production avec des parties textuelles stratégiques comme l'introduction et la conclusion. Dans l'introduction, nous relevons essentiellement les idées suivantes : le contexte de la recherche, le thème de recherche, l'implication personnelle (motivation ou origine de l'intérêt pour le sujet), l'annonce des questions de départ et la problématique, le choix méthodologique et l'annonce du plan (Tran, 2010). L'enseignant peut laisser travailler les apprenants sur un exemple d'introduction et les laisser relever les idées essentielles dans le texte. L'enseignant rassemble ensuite les idées avec les apprenants et systématise les points principaux. À l'étape de production, l'enseignant peut les faire rédiger une introduction sur

papier ou à l'aide de ScribPlus⁵¹. Il s'agit d'un outil TICE qui propose une aide logicielle pour la rédaction auprès des étudiants étrangers. L'enseignant peut s'en inspirer pour concevoir des séquences didactiques qui correspondent aux objectifs et au niveau des apprenants. L'activité de compréhension écrite se déroule avec une approche inductive. Le fait de repérer les structures discursives essentielles dans l'introduction apporte aux apprenants non seulement des connaissances sur la structuration de cette partie textuelle, mais les aide également à mieux observer d'autres textes. De plus, pour les activités de production écrite, ils peuvent réutiliser ces connaissances discursives et structurales dans leur écrit. Nous insistons sur le fait que le repérage des idées principales de chaque partie stratégique permet aux apprenants de mieux se concentrer sur leur écrit. Ils savent également ce qu'ils doivent aborder dans l'introduction par exemple. Il en est de même pour les activités sur les conclusions (Tran, *ibid.*). Les apprenants peuvent trouver les idées suivantes dans une conclusion : annoncer ce qu'on va trouver dans la conclusion (reformuler les questions de départ ou la problématique), la méthodologie, l'évaluation, les résultats, les limites et les perspectives. Nous avons introduit un exemple de repérage des structures discursives dans l'introduction (Cf. **Fiche pédagogique 1**) et dans la conclusion (Cf. **Fiche pédagogique 3**).

En outre, l'enseignant peut également recourir aux fonctions rhétoriques que nous avons évoquées précédemment pour développer la stratégie de lecture *top-down* ; celles-ci s'avèrent indispensables pour les écrits scientifiques. Le fait de sensibiliser les apprenants à ces fonctions rhétoriques dans un premier temps va les aider à prendre conscience des spécificités des écrits scientifiques et de l'importance à les réinvestir dans leur écrit. Pour les aider dans cette tâche, un outil, Dicorpus, a été développé au Lidilem.

➤ **Dicorpus, dictionnaire basé sur corpus**

Nous abordons dans cette partie un outil de travail qui cherche à tirer profit des descriptions linguistiques pour une utilisation didactique. Nous partons d'un constat mentionné par Cavalla & Loiseau (2013) concernant la nécessité de mettre en place un outil qui allie le corpus et les séquences didactiques. Selon eux, « les didacticiens ne peuvent pas toujours utiliser directement les outils développés par les linguistes dans le cadre de la linguistique de corpus » (*ibid.* : 179), car ils ne sont pas linguistes. De plus, comme nous

⁵¹ Scribplus est un logiciel en ligne et accessible à l'adresse suivante : <http://scribplus.ovh.org/>. Il a été conçu par F. Mangenot, Lidilem, Université Grenoble Alpes et A. Khara, EIAL, Université Stendhal-Grenoble 3.

l'avons également évoqué plus haut, les apprenants semblent apprécier les exemples présélectionnés par l'enseignant. L'outil que nous souhaitons mettre en place constitue une passerelle entre le corpus Scientext⁵² et les activités didactiques tout en nous appuyant sur les descriptions linguistiques du chapitre 5. L'outil de travail que nous allons présenter peut être considéré comme la première étape pour une transposition didactique. Parmi les outils en la matière, le principe de l'*Interactive LanguageToolbox*⁵³ nous intéresse particulièrement. Ce site, qui fonctionne comme un moteur de recherche, bénéficie de nombreuses ressources lexicographiques disponibles sur le web pour mettre à disposition des utilisateurs (enseignants, apprenants, etc.) un moyen de travail de différentes fonctionnalités (Verlinde, 2013).

Dans notre travail, afin d'aider l'enseignant à élaborer son parcours pédagogique et à sélectionner les exemples pertinents pour l'enseignement, nous nous inspirons du modèle développé par Tutin & Falaise (2013)⁵⁴ pour mettre à disposition de l'enseignant et des apprenants une base de données des séquences lexicalisées à fonction discursive (Figure 6-3). Cette base de données fonctionne comme un dictionnaire basé sur un corpus ou « dictionary-cum-corpus ». Elle allie les avantages de la linguistique de corpus et les techniques du TAL pour une aide à la rédaction scientifique pour les étudiants FLE (Tutin & Falaise, *ibid.*). Comme notre corpus SHS est annoté syntaxiquement et analysé linguistiquement, cela permet l'extraction de séquences lexicalisées à fonction discursive et la sélection des exemples les plus intéressants pour l'enseignement. Cette dernière étape est très importante, car les requêtes sur le corpus peuvent générer des exemples non pertinents que les apprenants ne sont pas encore en mesure d'enlever.

Le module conçu par Falaise, et appelé Dicorpus⁵⁵, permet de combiner les résultats de requêtes sur Scientext avec une base préalablement élaborée par l'enseignant. Notre base de données des séquences lexicalisées à fonction discursive s'appuie sur la typologie des séquences lexicalisées à fonction discursive que nous avons présentée dans le chapitre 4 et sur les descriptions linguistiques faites dans le chapitre 5. Les apprenants auront une entrée onomasiologique et une entrée sémasiologique. Dans la Figure 6-3, il s'agit de l'accès

⁵² En raison du droit de l'auteur, une partie de notre corpus SHS est accessible uniquement en Intranet. Il est à noter que pour nous utilisons uniquement dans Dicorpus des exemples extraits du corpus de libre de droit de Scientext.

⁵³ Site : <https://ilt.kuleuven.be/inlato/>

⁵⁴ Tutin & Falaise (2013) conçoivent une base de données pour traiter des séquences lexicalisées à fonction modale.

⁵⁵ Nous tenons à remercier Achille Falaise de son aide généreuse dans la conception de cette base de données.

onomasiologique dans Dicorpus. L'enseignant et les apprenants peuvent accéder à une structure discursive correspondant à leur recherche (reformulation, explication, etc.). Pour chaque MD, ils ont à leur disposition des descriptions des caractéristiques syntaxiques et sémantiques des MD (la fréquence dans le corpus, la portée, la position dans la phrase ainsi qu'une définition brève).

Séquences lexicalisées à fonction discursive dans les écrits scientifiques			
MODE D'ACCÈS ► Accès par sens ► Accès par expression	POUR EXPRIMER... comparaison ... concession ... condition ... exemplification ... causalité ... introduction thématique ... finalité ... opposition ... reformulation	EXPRESSION en somme (*) c'est-à-dire (***) à savoir (***) autrement dit (***) en d'autres termes (*) en résumé (*) en un mot (*) au total (**) somme toute (*) en définitive (*) en fin de compte (*) au fond (*)	SOUS-CATÉGORIE clarification POSITION DANS LA PHRASE postverbale tête de phrase PORTÉE locale DÉFINITION s'emploie pour introduire une explication, une définition, une interprétation afin de rendre l'énoncé plus clair

Figure 6-3 : Base de données des séquences lexicalisées à fonction discursive dans Dicorpus – Entrée onomasiologique

Il s'agit dans la Figure 6-4 d'une entrée sémasiologique de Dicorpus, c'est-à-dire une entrée par MD. Nous allons décrire plus en détail cette approche dans la partie suivante. À l'instar de l'approche onomasiologique, l'utilisateur peut accéder à toutes les données du marqueur.

MODE D'ACCÈS ► Accès par sens ► Accès par expression	EXPRESSION en somme en somme (*) c'est-à-dire (***) à savoir (***) autrement dit (***) en d'autres termes (*) en résumé (*) en un mot (*) au total (**) somme toute (*) en définitive (*) en fin de compte (*) au fond (*)	SOUS-CATÉGORIE récapitulation POSITION DANS LA PHRASE postverbale tête de phrase PORTÉE globale (proposition) DÉFINITION introduit une reformulation qui consiste en une synthèse d'idées. Ces idées sont concordantes
---	--	---

Figure 6-4 : Dicorpus - Entrée sémasiologique

En accédant par l'approche onomasiologique ou par l'approche sémasiologique, en choisissant chaque MD, l'utilisateur dispose des exemples présélectionnés et sauvegardés à

partir des requêtes sur le corpus Scientext. L'enseignant peut utiliser ce concordancier pour faire travailler les apprenants. Dans la Figure 6-5, figurent des exemples sauvegardés pour *en somme* et intégrés à l'interface de Dicorpus. L'utilisateur a la possibilité d'afficher le contexte large de chaque occurrence (Figure 6-6) s'il le souhaite.

Exemples validés sur le corpus Scientext			
▼ N°	▼ Contexte gauche: 50 mots	Occurrence:	▼ Contexte droit: 50 mots
1		En somme	, nos connaissances suggèrent désormais de remplacer l'opposition traditionnelle génétique/épigénétique (Figure 1A) par une vision graduelle reposant sur quatre échelles corrélées de changements, chimique, spatiale et temporelles (Figure 1B) .
2		En somme	, les périodicités observées sont compatibles avec une épi organisation solénoïdale du chromosome, qui regrouperait dynamiquement en des points focaux les deux partenaires transcriptionnels, les facteurs de transcription et leurs sites de liaison sur l'ADN (Figure 2B) .
3	En optant pour un référentiel autre,	en somme	une épistémé différente, pour légitimer une nouvelle définition de soi, les personnes étiquetées peuvent échapper à la confrontation .
4	Le candidat cherche	en somme	à suivre une fabrique artificielle du sens selon les ficelles d'une mise en scène convenue (Goffman 1973) en prenant en compte les contraintes institutionnelles telles qu'il les appréhende, et les lois qu'il suppose être celles du genre littéraire attendu .
5		En somme	, le projet pédagogique, dans ses dimensions fondamentales, est un objet de réflexion partagé par les enfants et les adultes .
6		En somme	, le tuteur, tel un formateur APP, assure le soutien cognitif en jouant le rôle de facilitateur plutôt que de dispensateur de savoir .
7		En somme	, comme à l'APP, le tuteur / formateur s'efforce de donner à l'apprenant les moyens de se repérer dans sa progression et de prendre conscience des processus qu'il utilise pour apprendre .
8	Ainsi Paul Valéry rappelle dans " Variétés " que " Tout sport sérieusement pratiqué exige des épreuves, des privations parfois sévères, une hygiène, une tension et une constance mesurables par les résultats,	en somme	une véritable morale de l'action " (Valéry, 1957) .

Figure 6-5 : Dicorpus - Exemples validés sur le corpus Scientext

Exemples validés sur le corpus Scientext	
Le candidat cherche	en somme
5	En somme
[sed-com-texte_communication_scientext_205-body] - La Fabrique de l'Institution: Dynamiques d'ajustement du sens et inscription symbolique dans l'espace éducatif (Actes AREF 2007)	
: ainsi les ceintures de comportement, les responsabilités confiées aux enfants, le parrainage (une forme de tutorat), les espaces de parole[7] .	
1.2. La dimension symbolique à l'ordre du jour	
La Pédagogie Institutionnelle revendique la prise en compte de " l'inconscient " en jeu dans le groupe, comme nouvelle dimension donnée aux pratiques de la classe coopérative développées par Freinet. Cette réflexion est poursuivie par l'équipe de l'EE, qui s'attache à mettre en scène la dimension symbolique au coeur même des espaces d'échange vécus par les enfants. Lorsque je parle de l'EE comme d'un espace symbolique, il ne faut pas simplement entendre que cette dimension fait partie de la réflexion de l'équipe adulte. Il s'agit surtout de souligner le fait que la question du " sens " est mise à l'ordre du jour des réunions d'enfants. L'une des particularités du projet est ainsi que le groupe lui-même est amené à s'interroger sur le sens des différentes institutions et sur leurs fonctions, sur la signification des mots qu'il emploie, sur les origines des énoncés rituels, etc. En somme, le projet pédagogique, dans ses dimensions fondamentales, est un objet de réflexion partagé par les enfants et les adultes.	
1.3. La question des dynamiques d'inscription	
La thèse en cours a pour objet l'analyse des dynamiques selon lesquelles les enfants s'inscrivent dans	

Figure 6-6 : Dicorpus - Contexte large de *en somme*

La particularité de cette base de données est qu'elle combine les ressources du corpus Scientext ainsi que les descriptions linguistiques effectuées précédemment. Elle permet à l'enseignant et à l'apprenant d'accéder directement aux exemples sélectionnés par l'enseignant sans passer par les requêtes sur le corpus et ils peuvent ainsi se concentrer sur le phénomène linguistique qui les intéresse. L'enseignant peut s'inspirer de ces exemples authentiques pour développer différentes activités.

Nous proposons d'ajouter une autre option à Dicorpus afin de mieux répondre aux différents besoins de l'utilisateur. Comme nous l'avons montré dans le chapitre 5, la *portée* est une notion intéressante à aborder pour une analyse des MD. L'analyse de la portée du MD sert à révéler son fonctionnement syntaxique et sémantique. Il est important de mettre sur Dicorpus toutes ces informations. Une information générale sur la *portée* comme la portée globale ou la portée locale, pourrait ne pas satisfaire l'utilisateur qui souhaite consulter le fonctionnement général d'un MD. Nous proposons donc d'ajouter deux niveaux d'utilisation pour la *portée* et pour les exemples sauvegardés sur le corpus Scientext. Pour la *portée*, au premier niveau, l'utilisateur n'a que des informations sur la *portée globale* ou la *portée locale* (il s'agit de la fonctionnalité dont dispose Dicorpus actuellement, Cf. Figure 6-3 et Figure 6-4). Au deuxième niveau, nous proposons d'ajouter toutes les informations sur la portée des MD que nous avons mises dans les Tableaux récapitulatifs des caractéristiques des MD au chapitre 5. L'utilisateur aura donc les informations détaillées et savoir si les MD peuvent porter sur une proposition, sur un syntagme verbal, sur un syntagme nominal, etc. Toutes ces informations sont très intéressantes pour l'enseignement/apprentissage des MD, car comme nous l'avons vu précédemment, ces propriétés syntaxiques constituent également un indice discriminant. Il en est de même pour les exemples. Nous proposons d'ajouter le premier niveau d'accès où l'utilisateur a à sa disposition des exemples représentatifs de chaque MD, ce qui s'avère utile pour l'utilisateur qui veut visualiser rapidement le fonctionnement des MD. Au deuxième niveau, l'utilisateur accèdera aux informations plus précises et à des analyses plus approfondies, comme le proposent les fonctionnalités actuelles de Dicorpus (Figure 6-6).

Comme nous venons de le voir, tant l'exemple de stratégie de Lundquist que le développement de Dicorpus font appel à deux entrées didactiques complémentaires déjà évoquées : onomasiologie et inductif complétées par sémasiologie et déductif. Nous nous proposons de décrire plus précisément ce dernier couple afin que nos fiches pédagogiques puissent apparaître complètes à tous les niveaux didactiques et méthodologiques.

6.3.1.2. Approche déductive et approche sémasiologique

L'approche inductive présente de nombreux avantages, mais comme nous l'avons vu, les apprenants ont parfois du mal à formuler les règles et l'apprentissage peut être très long. L'enseignant doit intervenir pour expliciter certaines notions ; nous proposons dans ce cas d'utiliser l'approche déductive. De plus, même si l'approche onomasiologique s'avère indispensable pour la production écrite, nous n'allons pas exclure l'approche sémasiologique

dans l'enseignement/apprentissage des MD. En effet, une entrée par unité lexicale est aussi intéressante notamment pour des MD un peu plus difficiles à comprendre. Si les approches inductives et onomasiologiques sont très importantes pour la construction des connaissances et la conceptualisation des règles, les approches déductive et sémasiologique permettent de consolider les connaissances. Elles servent notamment à développer les stratégies de lecture *bottom-up* proposées par Lundquist (*op.cit.*). Nous allons traiter dans cette partie du rôle que jouent les approches déductive et sémasiologique dans l'enseignement/apprentissage des MD et le développement des stratégies de lecture *bottom-up*.

a. Approche déductive

Le type de raisonnement qu'exprime l'approche déductive va des règles aux exemples, du général vers le particulier. Elle complète l'approche inductive. D'une part, cette approche permet aux apprenants d'accéder à d'autres exemples et de consolider les connaissances obtenues à l'issue de l'approche inductive. D'autre part, si les apprenants ont du mal à formuler les règles, l'enseignant peut intervenir en explicitant les règles et fournir d'autres exemples pour les vérifier. Nous sommes d'accord avec Boulton & Tyne (2014) quand ils évoquent la place d'une entrée déductive pour l'apprentissage : « le travail sur corpus peut tout aussi bien se faire de façon déductive, ce qui permettrait plus facilement la réutilisation de matériels généralisables » (*ibid.* : 166).

Avant de commencer un cours sur la reformulation, l'enseignant peut faire travailler les apprenants sur différentes formes de reformulation (reformulation par paraphrase, ou avec des marqueurs de reformulation, etc.), et précise que le cours aborde uniquement la reformulation effectuée par des moyens lexicaux. L'enseignant peut par la suite montrer l'utilisation des marqueurs de reformulation en lien avec des fonctions comme clarifier, récapituler, reconsidérer. Cependant, il n'est pas recommandé de présenter la liste complète des marqueurs de reformulation : cela n'a pas d'intérêt sur les plans didactique et cognitif puisque les apprenants ne sauront pas forcément où et comment les utiliser.

L'enseignant peut utiliser les approches inductive et déductive en alternance. Nous proposons de commencer l'apprentissage par une entrée inductive suivie par une approche déductive. En effet, si l'enseignant démarre le programme par l'approche inductive, c'est-à-dire qu'il laisse les apprenants observer le corpus et les diverses formes linguistiques et qu'ensuite il les aide à formuler la règle, cela paraît souvent plus pertinent selon les études

citées précédemment. Ainsi, l'enseignant permet aux apprenants de 1/ développer des compétences cognitives et 2/ relever les régularités de fonctionnement syntaxique et sémantique des MD que les apprenants auront détectées à l'issue de l'approche inductive serviront à éclairer un travail par l'approche déductive. C'est pourquoi l'entrée par déduction représente une étape de mémorisation des connaissances en mettant à disposition des apprenants de plus amples exemples.

Si la recherche des règles d'usage de *à savoir* semble plus facile, ce n'est pas le cas pour les MD de reconsidération, comme *en définitive* ou *en fin de compte* par exemple (Cf. Exercices pour les marqueurs de reformulation *en définitive/en fin de compte*, **Fiche pédagogique 2**, page 258). En cas de difficulté, l'enseignant peut intervenir pour aider les apprenants à relever les propriétés de ces marqueurs. Nous pouvons donc conclure cette partie en affirmant que l'approche inductive et l'approche déductive sont complémentaires et très propices à l'utilisation du corpus pour l'enseignement des phénomènes linguistiques.

b. Approche sémasiologique

L'approche sémasiologique décrit une démarche inverse par rapport à l'approche onomasiologique. Elle complète l'approche onomasiologique et peut être utile à l'enseignement/apprentissage du lexique en général. L'apprentissage par l'approche sémasiologique, qui va de la forme au sens, est très fréquente dans les méthodes de FLE. Nous ne préconisons pas un enseignement par listes de mots, car cela ne présente pas beaucoup d'intérêt a priori pour les apprenants. Pourtant, un accès par unité lexicale peut être parfois utile sur le plan cognitif et en compréhension, en particulier pour les étudiants de niveau intermédiaire. Cette approche accorde la place à l'imitation ou à la répétition, ce qui aide à la mémorisation des structures syntaxiques et des valeurs sémantiques. L'approche sémasiologique peut être suivie de l'approche déductive. Les apprenants travaillent d'abord sur un MD et ils peuvent ensuite recourir à d'autres exemples extraits du corpus par une entrée déductive pour valider ou infirmer une conclusion. Dans notre base de données Dicorpus (Figure 6-4), les apprenants peuvent aussi accéder directement aux MD par une entrée sémasiologique.

Pour des approches déductive et sémasiologique, l'enseignant peut utiliser des exercices lacunaires. Nous n'excluons pas le rôle de ce type d'exercice à certaines étapes de l'apprentissage, par exemple pour la vérification ou l'évaluation. Comme l'ont montré

Montagnon (2007) et Cavalla (à paraître), les exercices lacunaires peuvent « aider à la découverte et à la prise de conscience de phénomènes linguistiques en tant que structures figées (ou quasi figées) » (Cavalla, à paraître : 4). Ce genre d'exercices peut être introduit suite aux activités avec l'approche inductive et onomasiologique. La question qui se pose alors est de savoir comment les mettre en place pour éviter l'impression qu'on peut remplir les trous « au hasard ». L'enseignant peut proposer aux étudiants de niveau avancé des exercices lacunaires tout en leur demandant de préciser leur choix. Autrement dit, ils sont invités à expliquer pourquoi il faut utiliser tel ou tel MD et pourquoi deux MD ne sont pas interchangeables par exemple, tout en se basant sur des raisons syntaxiques ou sémantiques des MD qu'ils auront acquis antérieurement (Cf. 6.3.2 Exercices avec les concordanciers, page 241).

Pour conclure cette partie sur les différentes approches que nous venons de présenter, nous pensons que les enseignants doivent utiliser ces différentes approches avec souplesse dans l'enseignement des phénomènes linguistiques à l'aide du corpus. Si l'approche inductive est très adaptée à un travail sur corpus, l'approche onomasiologique met à disposition des apprenants une entrée par des fonctions discursives spécifiques des écrits scientifiques. L'approche déductive et l'approche sémasiologique sont susceptibles d'aider les apprenants à consolider leurs connaissances.

En élaborant un cours d'enseignement/apprentissage des MD, l'enseignant peut s'appuyer de prime abord sur notre typologie des séquences lexicalisées à fonction discursive pour élaborer le programme d'enseignement/apprentissage des MD. Ensuite, l'apprenant est amené à travailler sur le corpus, selon une approche inductive, pour relever les régularités syntaxiques et sémantiques du MD. En cas de difficultés pour la détection du fonctionnement d'une unité lexicale, l'enseignant peut intervenir pour donner des règles ou des explications (approche sémasiologique). Enfin, l'apprenant accède à d'autres exemples qui véhiculent la même valeur sémantique du MD (approche déductive) pour vérifier la règle qu'il aura relevée ou qui aura été explicitée par l'enseignant. Sachant que l'enseignant et les apprenants ont à leur disposition une base de données des séquences lexicalisées à fonction discursive, ils peuvent y accéder sans passer par des requêtes sur le corpus Scientext.

c. Stratégie de lecture *bottom-up*

Si la stratégie *top-down* s'appuie sur les connaissances du plan structurel des apprenants, la stratégie *bottom-up* fonctionne dans le sens inverse : elle propose une entrée sémasiologique par certains mots du texte et nous nous en servons de la sorte. Il s'agira de faire repérer les MD d'un texte et de voir l'enchaînement entre les MD, leur position entre eux et de voir ensuite les idées qu'ils introduisent ou rassemblent.

La reconnaissance des caractéristiques des MD permet à l'apprenant de repérer plus facilement les MD et de deviner les rôles/idées qui les entourent. Cette stratégie peut faciliter le repérage de ces structures discursives. Pour la compréhension écrite, en recourant à la segmentation en constituants, l'apprenant a plus de facilité pour relever les structures discursives. En fait, toutes les connaissances que l'apprenant aura acquises à la suite de l'apprentissage des MD serviront à mieux repérer le contenu que l'auteur souhaite transmettre dans l'énoncé. Par exemple, l'apprenant peut facilement repérer une succession d'idées développées grâce aux MD qui indiquent l'enchaînement thématique comme *premièrement/en premier lieu/tout d'abord, deuxièmement, enfin, etc.*, ou bien une introduction de la nouvelle thématique grâce aux marqueurs comme *en ce qui concerne, quant à, au sujet de, etc.* (Cf. **Fiche pédagogique 1**, page 249).

Les deux stratégies de lecture *top-down* et *bottom-up* se complètent et sont tout à fait adaptables à notre corpus d'écrits scientifiques. L'enseignant peut commencer d'abord par le *bottom-up* pour enseigner les MD – les unités servant à structurer le discours. Il peut par la suite sensibiliser les apprenants aux structures discursives assez régulières dans les écrits scientifiques. L'alternance de ces deux approches va profiter non seulement à la compréhension écrite, mais également à la production écrite des apprenants. Ceux-ci apprennent à construire au fur et à mesure des stratégies de production. En effet, toutes les activités aident les apprenants à se familiariser avec le lexique scientifique, les tournures scientifiques et les structures discursives. Le corpus représente un modèle pour l'organisation du texte et également une source de données à laquelle les apprenants peuvent recourir pour vérifier comment les locuteurs natifs utilisent telle ou telle tournure ou telle unité phraséologique. Ils y trouveront des moyens linguistiques, des modèles afin de les réinvestir dans leur écrit.

Nous nous inspirons des stratégies de lecture *top-down* et *bottom-up* dans les travaux de Lundquist (2013) pour proposer un schéma appliqué dans un écrit scientifique comme dans la Figure 6-7 : Schéma des stratégies de lecture top-down *et* bottom-up.

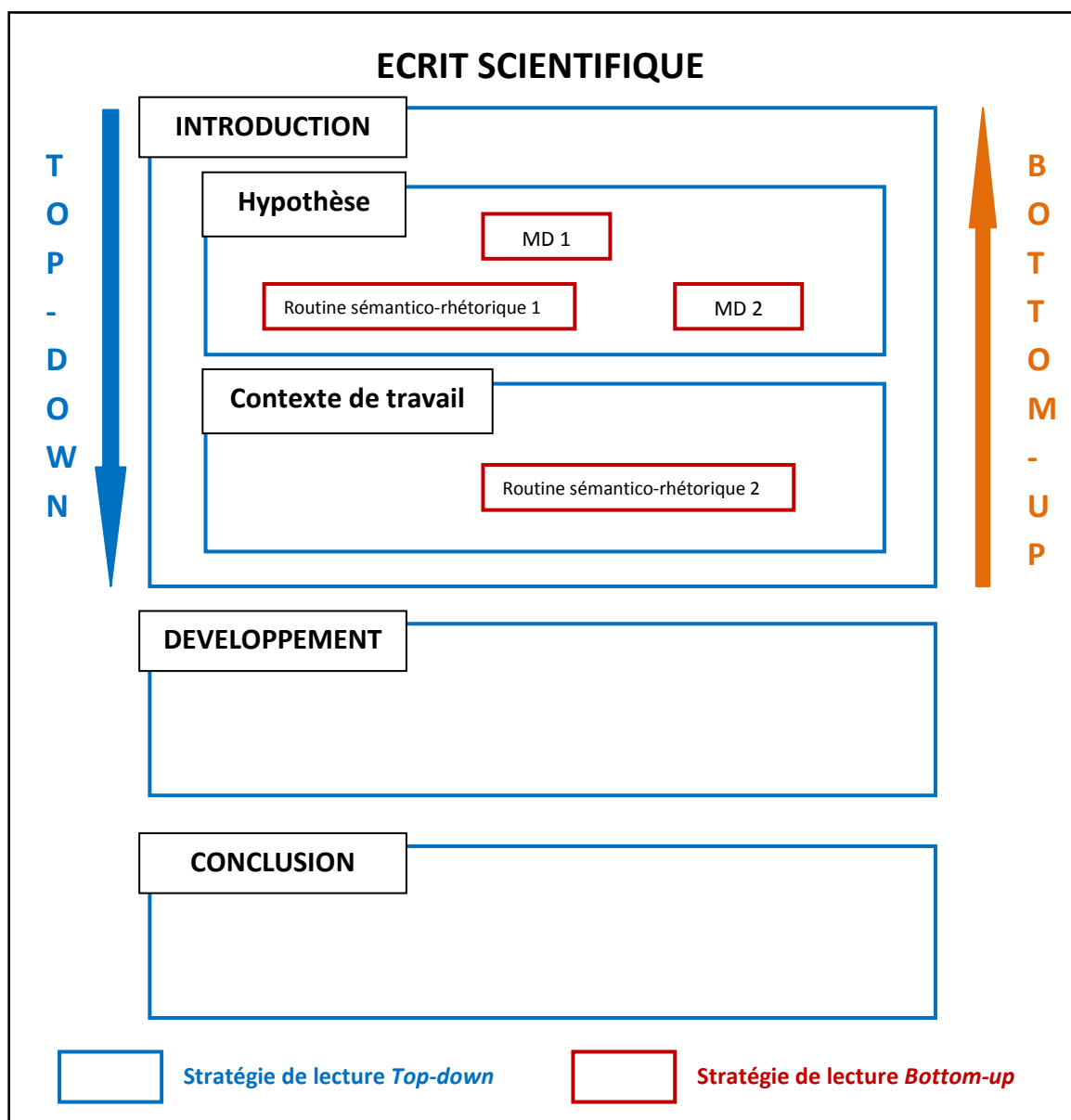


Figure 6-7 : Schéma des stratégies de lecture *top-down* et *bottom-up*

En résumé, ces stratégies *top-down* et *bottom-up* permettent d'éviter une lecture linéaire qui n'est pas toujours très efficace pour le traitement du texte. Ces stratégies s'avèrent indispensables aussi bien pour les non natifs que pour les locuteurs natifs. Si la stratégie *top-down* apporte à l'apprenant une vue d'ensemble sur le texte grâce aux connaissances sur les structures discursives, la stratégie *bottom-up* consiste en une lecture plus approfondie en analysant les constituants de l'énoncé (MD ou « routines sémantico-rhétoriques »). Si la stratégie *top-down* traite le texte au niveau macrotextuel (les structures rhétoriques et discursives), la stratégie *bottom-up* se trouve au niveau microtextuel (les relations syntagmatiques et les rapports sémantiques).

Pour récapituler, nous avons présenté dans cette partie quelques pistes d'enseignement/apprentissage des MD en associant plusieurs approches. Ces approches qui sont complémentaires peuvent être utilisées en alternance en classe de langue en fonction du niveau et de la progression de l'apprenant. L'association de différentes approches a pour objectif de développer chez les étudiants des compétences linguistiques et discursives voire des stratégies d'apprentissage. Ces stratégies sont indispensables pour aborder d'autres contenus universitaires. D'un côté, l'association des approches inductive, onomasiologique et les stratégies de lecture *top-down* nous semblent pertinentes pour la prise de conscience des phénomènes linguistiques en prenant appui sur un corpus. D'un autre côté, les approches déductive, sémasiologique et les stratégies de lecture *bottom-up* peuvent être regroupées pour la consolidation des connaissances lexicales acquises à la suite des activités précédentes. Les réflexions que nous avons développées peuvent être considérées comme des suggestions pour la conception d'une méthodologie générale pour l'enseignement d'un phénomène linguistique spécifique en classe de langue et de l'utilisation d'un corpus.

Il est important de rappeler que les MD ne sont qu'une partie de la chaîne qui conduit à l'écrit universitaire. Les propositions didactiques que nous allons aborder dans la partie suivante prennent les MD comme exemples, mais ces propositions peuvent être élargies à d'autres éléments linguistiques (collocations, « routines sémantico-rhétorique », etc.). L'initiation aux écrits scientifiques devrait bien entendu prendre en compte l'ensemble de ces éléments.

6.3.2. Propositions didactiques

Nous allons présenter quelques fiches pédagogiques élaborées après ces réflexions didactiques afin de tenter de mettre en pratique les différentes approches évoquées. L'enseignant y trouvera quelques suggestions pour l'enseignement/apprentissage des MD ainsi que des activités qu'il pourra développer en classe de langue. Comme nous souhaitons montrer à l'enseignant différents aspects à traiter pour un écrit scientifique parmi lesquels les MD, les « routines sémantico-rhétoriques », nous avons décidé de faire travailler les apprenants sur des extraits plus ou moins longs d'un écrit scientifique. Nous distinguons la dimension phrastique de la dimension textuelle. En fait, si le choix des exemples relativement courts venant de différentes disciplines n'est pas une chose facile, le choix des textes complets est encore plus complexe. Au niveau de la phrase, le problème s'arrête pratiquement

au plan terminologique, l'enseignant peut faire travailler les apprenants sur les exemples relativement courts venant de différentes disciplines pour faire repérer les propriétés syntaxiques et sémantiques d'un élément linguistique. Ce n'est pas l'unique problème quand les apprenants doivent travailler sur un texte plus long, un article scientifique par exemple, car en plus du savoir architectural et linguistique (Cavalla, 2007), il est demandé aux apprenants le savoir disciplinaire pour comprendre le texte. L'enseignement/apprentissage des MD sera bénéfique uniquement quand les apprenants comprendront le texte, ce qui est difficile quand on n'est pas spécialiste d'une discipline. Si l'accès au sens du texte doit être un vrai travail de déchiffrement, il est inutile d'envisager l'accès au sens des MD pour les utiliser à bon escient. Il n'est donc pas conseillé de leur donner de tels textes. Par exemple, pour un accès au sens facilité des MD (et autres phénomènes linguistiques), il vaut mieux éviter que des étudiants en linguistique travaillent sur un texte en médecine et vice versa. Notons qu'il existe un mouvement de FOU par discipline avec l'édition d'ouvrages spécialisés pour les apprenants, par exemple la méthodologie pour les ingénieurs de Carras & al. (2014).

C'est pourquoi les exemples-textes que nous avons choisis pour les activités didactiques s'adressent essentiellement aux étudiants de la discipline en question. L'enseignant peut utiliser ces exemples pour les étudiants d'une autre discipline, mais cela ne sera pas évident.

Avant de proposer des exercices concrets et des fiches pédagogiques, nous souhaitons ouvrir une parenthèse sur les différentes tâches que l'enseignant peut envisager avant un cours d'enseignement/apprentissage des phénomènes linguistiques basé sur corpus.. Nous considérons cette étape comme une étape de préparation ou d'échauffement pour l'enseignant ou pour les apprenants. L'enseignant trouvera ici quelques tâches à préparer avant de démarrer un cours d'enseignement avec un corpus. Pour les enseignants qui n'ont pas encore travaillé sur corpus, l'intégration du corpus en classe de langue leur demande évidemment de s'adapter à une pratique d'enseignement qui recourt à un nouvel outil pour enseigner des phénomènes linguistiques. Avant d'élaborer son cours à l'aide du corpus, l'enseignant doit prendre conscience que l'utilisation du corpus peut privilégier un apprentissage en autonomie des apprenants si l'on envisage de les faire travailler directement sur les corpus.

En outre, pour les apprenants n'ayant jamais travaillé sur le corpus, l'enseignant doit expliquer les aspects techniques : comment effectuer des requêtes, aider les apprenants à

apprivoiser l'interface du corpus. Notre corpus SHS (intégré au corpus Scientext⁵⁶) est doté d'une interface conviviale et facile à utiliser, même pour les non-linguistes. Trois modes de recherches sont mis en place (recherche sémantique, recherche libre, recherche avancée). Les apprenants peuvent commencer par des recherches sémantiques qui sont prêtes à l'emploi ou effectuer quelques recherches simples.

Nous avons insisté à plusieurs reprises sur le rôle de l'enseignant dans l'organisation du programme d'apprentissage, surtout dans la sélection des exemples qui sont faciles à comprendre pour des étudiants. Afin d'aider l'enseignant à préparer des parcours pédagogiques destinés à l'enseignement des marqueurs de reformulation, nous envisageons de mettre en place un tableau détaillé des valeurs sémantiques de ces marqueurs. Ce tableau résulte des descriptions linguistiques dans le chapitre 5 et d'une présélection de notre part. Comme nous l'avons vu dans le chapitre 5, chaque MD peut avoir différentes valeurs sémantiques. Pour faciliter la compréhension, toutes les occurrences ont été classées selon ce critère. L'enseignant y trouvera des exemples pour chaque valeur sémantique et peut s'en inspirer pour développer des activités pédagogiques. Dans l'explication des phénomènes linguistiques, nous proposons à l'enseignant d'éviter les terminologies et le métalangage, ce qui risque d'augmenter la surcharge cognitive des apprenants, voire ne pas leur être utile du tout. Par exemple, les termes que nous utilisons dans l'analyse linguistique ne peuvent pas être employés tels quels en classe de langue. Il est nécessaire de procéder avant tout à la didactisation de la terminologie.

Comme nous l'avons évoqué précédemment, l'enseignant peut faire travailler les apprenants sur des extractions du concordancier. Nous souhaitons présenter quelques exemples-exercices pour montrer comment enseigner les MD avec un concordancier.

6.3.2.1. *Exercices d'entraînement*

Nous présentons dans cette partie trois exemples-exercices sur l'utilisation du concordancier dans l'enseignement/apprentissage. Il s'agit des exercices sur *c'est-à-dire*, à *savoir* et sur la distinction de deux marqueurs.

⁵⁶ Lien vers Scientext : <http://scientext.msh-alpes.fr/scientext14/?tab=texts-criteria>

Exercice 1 : Concordancier de *c'est-à-dire*

Les phrases suivantes qui sont extraites du concordancier de *c'est-à-dire* du corpus Scientext.

1. À condition toutefois de désenclaver la réception, c'est-à-dire de définir celle-ci comme une activité socialisée et nomade, indépendamment du caractère éventuellement résidentiel des objets techniques. [infcom-art-5-body]
2. Il s'ensuit que les séries télévisées constituent des vecteurs majeurs pour connaître la police ; il s'agit d'un savoir collectif acquis par procuration, c'est-à-dire par médias interposés. [infcom-art-5-body]
3. Ce qui explique, par exemple, que la principale forme de judiciarisation puisse être le « judicialreview » c'est-à-dire le contrôle des lois par les tribunaux (Vallinder, 1994 ; Hertog et Halliday, 2004) ou, plus précisément encore, le « constitutionalreview » (Shapiro et Stone, 2004). [soc-art-35-body]
4. D'une part, c'est la seule qui marque sa singularité en s'appuyant sur les conduites motrices des élèves ; d'autre part, c'est également la seule qui est confondue avec les supports qu'elle emploie, c'est-à-dire avec les activités physiques et sportives. [sed-art-52-body]
5. Les pédagogues, c'est-à-dire ceux qui vont inscrire leur vie principalement autour des questions éducatives, ne vont cesser de s'y référer. [sed-art-383-body]
6. Comme d'autres revues avant elle, ACP fonctionne selon le modèle économique dit « auteur payeur », c'est-à-dire que la gratuité de consultation est financée par le paiement assuré par l'auteur (ou plus généralement par son institution) au moment où le manuscrit est rendu accessible dans la partie discussion de la revue. [infcom-art-432-body]
7. Tous les échanges, horizontaux et verticaux, se font de manière asynchrone, c'est-à-dire que les participants et le modérateur animateur se connectent à l'heure de leur choix, sans obligation de simultanéité à des moments particuliers. [infcom-art-455-body]
8. La troisième ligne est une analyse intra- bancaire (within), c'est-à-dire que l'on travaille en écart par rapport aux moyennes par banques. [eco-art-572-body]

Observez les phrases suivantes qui sont extraites du concordancier de **c'est-à-dire** du corpus Scientext et répondez aux questions ci-dessous.

- Observez le contexte gauche et le contexte de droite de **c'est-à-dire**. Quelle est leur nature syntaxique ?
- Définissez la relation entre le contexte gauche et le contexte de droite de **c'est-à-dire**. Remplissez le tableau suivant par les indices contextuels que vous allez relever dans cet exercice.
- D'après vous, quel est le rôle de **c'est-à-dire** dans ces phrases ? Quelle est la valeur qu'il véhicule ?

Phrase	Contexte gauche	Contexte de droite	Rapport entre 2 contextes
1.			
2.			
...			

Approche : inductive guidée + onomasiologique, aspect sémantico-sémantique.

Explication : *C'est-à-dire* fait partie des MD les plus productifs de notre corpus (704 occurrences). Comme nous ne pouvons pas faire travailler les apprenants sur toutes les occurrences de ce marqueur, pour faire découvrir ses différentes valeurs, l'enseignant sélectionne préalablement quelques occurrences où *c'est-à-dire* a la même valeur d'usage (il s'agit ici de l'approche onomasiologique). Dans cet exercice, *c'est-à-dire* introduit une valeur métalinguistique (Cf. Tableau 5.1). Cette concordance peut être utilisée comme un exercice d'observation à l'appui de l'approche inductive avec une entrée onomasiologique pour la valeur préférentielle, mais sémasiologique pour le choix du MD spécifique (et pas tous les MD de la valeur sémantique). L'enseignant fait repérer le sens de *c'est-à-dire* dans ces exemples en demandant aux apprenants d'analyser la structure syntaxique du contexte gauche et du contexte droit de *c'est-à-dire*, ainsi que de relever le rapport entre les deux contextes et formuler le rôle que joue le MD.

L'enseignant fait remplir le tableau ci-dessus qui recensera les informations sur les propriétés syntaxiques et sémantiques du contexte gauche et du contexte de droite de *c'est-à-dire*. Il peut intervenir pour guider les apprenants à repérer ces propriétés. L'objectif du tableau est d'aider les apprenants à découvrir une certaine équivalence de la nature syntaxique entre les deux séquences établies par *c'est-à-dire*. L'enseignant se sert du Récapitulatif des propriétés syntaxiques et sémantiques de *c'est-à-dire* (Cf. Tableau 5.1) afin d'obtenir les idées essentielles à aborder pour les apprenants. *C'est-à-dire* peut porter sur un groupe nominal, un groupe verbal, un groupe prépositionnel, ou une phrase. Quand *c'est-à-dire* porte sur une proposition, *c'est-à-dire* est suivi de *que*. L'apprentissage évolue dans une approche inductive guidée.

Exercice 2 : Concordancier de à savoir

Les phrases suivantes sont extraites du concordancier de *à savoir* du corpus Scientext.

1. Plus précisément, il entend décrire et expliquer la réception des séries policières par les policiers réels, à savoir « ce qu'ils en font » et « ce que ça leur fait ». [infcom-art-5-body]
2. C'est aussi le cas des deux types d'objets circulant au sein du cercle de la kula, à savoir les colliers de petits coquillages rouges et les brassards de gros coquillages blancs. [ant-art-57-body]
3. D'une part, je vais illustrer à travers l'analyse de constructions diaphoniques comment s'articulent les trois composantes constitutives de tout discours, à savoir les dimensions

linguistique, textuelle et situationnelle. [lin-art-99-introduction]
4. Il permettait également d'associer les deux noyaux historiques de la ville, à savoir l'implantation portuaire orientale (parallèle à la Breedstraat) et la partie longtemps demeurée agricole à l'ouest. [his-art-363-body]
5. La pression sur les prix concerne surtout les deux grands produits des cabinets dans les années 1990 à savoir les projets de reengineering et la mise en place d'outils informatiques. [soc-art-56-notes]
6. Nous proposons d'aborder une problématique centrale de la géographie sociale et culturelle française, à savoir la production de l'espace social, en affirmant que ce dernier est également constitué par des discours. [geo-art-282-body]

Observez les phrases suivantes qui sont extraites du concordancier de **à savoir** et répondez les questions suivantes.

- Repérez le contexte de droite de **à savoir**. Quelle est sa nature syntaxique ?
- Repérez maintenant le contexte gauche de **à savoir**. Quelle est sa nature syntaxique ?
- Quel est le rapport entre le contexte gauche et le contexte de droite du marqueur ?

Notez dans le tableau ci-dessous les indices contextuels et leur nature syntaxique dans le tableau ci-dessous.

- Quelle est votre remarque sur le rôle de **à savoir** dans cet exercice ?

Phrase	Contexte gauche	Contexte de droite	Rapport entre 2 contextes
1.			
2.			
...			

Approche : inductive guidée + onomasiologique + déductive, aspect syntactico-sémantique.

Explication : L'activité a pour objectif de faire repérer la valeur énumérative, valeur significative de *à savoir*. L'enseignant guide les apprenants afin qu'ils relèvent eux-mêmes les propriétés du marqueur dans cet exercice en répondant au fur et à mesure aux questions sur les propriétés syntaxiques et sémantiques de *à savoir*. Les apprenants arrivent à formuler la règle de fonctionnement du marqueur : *à savoir* peut introduire une énumération ou une spécification. Cet exercice peut être suivi d'un autre exercice de reformulation que l'enseignant peut créer à partir d'autres exemples du concordancier *à savoir*. L'objectif de ce deuxième exercice est d'aider à la mémorisation. L'apprentissage se déroule dans ce cas dans une approche déductive et une approche sémasiologique.

Exercice 3 : Comparaison *c'est-à-dire* et *à savoir*

L'enseignant se sert des tableaux récapitulatifs des caractéristiques de *c'est-à-dire* (Cf. Tableau 5.1) et *à savoir* (Cf. Tableau 5.2) et le tableau de comparaison de ces marqueurs (Cf. Tableau 5.5) pour l'aider dans la conception des activités et pour le repérage des spécificités de chaque MD. Les exemples suivants sont extraits des concordanciers de *c'est-à-dire* et *à savoir*. L'enseignant enlève les marqueurs et laisse remplir les phrases aux apprenants.

Exemples originaux :

1. Dans les zones non loties (dénommées parfois habitat spontané), *c'est-à-dire* construites sans autorisation, le territoire scolaire peut être considéré comme illégal et l'école construite par les populations ou par des associations peut être détruite, à l'instar des habitations, sur ordre de l'administration. [ant-art-430-body]
2. Le concept de profession particulièrement médiatisée recouvre aussi un paradoxe, *à savoir* le couplage entre une omniprésence médiatique et une distance vis-à-vis de la société. [infcom-art-5-body]
3. Il permettait également d'associer les deux noyaux historiques de la ville, *à savoir* l'implantation portuaire orientale (parallèle à la Breedstraat) et la partie longtemps demeurée agricole à l'ouest. [his-art-363-body]
4. Tous les échanges, horizontaux et verticaux, se font de manière asynchrone, *c'est-à-dire* que les participants et le modérateur animateur se connectent à l'heure de leur choix, sans obligation de simultanéité à des moments particuliers. [infcom-art-455-body]
5. Ce travail très dynamique permet d'aboutir à la bonne formulation de patrons de fouille pour des relations générales qui soient efficaces sur ce corpus, *c'est-à-dire*, ramenant peu de mauvais exemples (bruit réduit) tout en étant le plus exhaustif possible (silence réduit). [lin-art-97-body]
6. Les mêmes musiques que celles utilisées dans l'étude princeps étaient reprises dans notre étude, *à savoir* un extrait provenant de l'Allegro de Mozart (Petite Musique de nuit, Divertimento). [psy-art-246-body]

Exercice :

Phrases	Explication
1. Dans les zones non loties (dénommées parfois habitat spontané),..... construites sans autorisation, le territoire scolaire peut être considéré comme illégal et l'école construite par les populations ou par des associations peut être détruite, à l'instar des habitations, sur ordre de l'administration. [ant-art-430-body]	

2. Le concept de profession particulièrement médiatisée recouvre aussi un paradoxe,..... le couplage entre une omniprésence médiatique et une distance vis-à-vis de la société. [infcom-art-5-body]	
3. Il permettait également d'associer les deux noyaux historiques de la ville,..... l'implantation portuaire orientale (parallèle à la Breedstraat) et la partie longtemps demeurée agricole à l'ouest. [his-art-363-body]	
4. Tous les échanges, horizontaux et verticaux, se font de manière asynchrone,..... les participants et le modérateur animateur se connectent à l'heure de leur choix, sans obligation de simultanéité à des moments particuliers. [infcom-art-455-body]	
5. Ce travail très dynamique permet d'aboutir à la bonne formulation de patrons de fouille pour des relations générales qui soient efficaces sur ce corpus,....., ramenant peu de mauvais exemples (bruit réduit) tout en étant le plus exhaustif possible (silence réduit). [lin-art-97-body]	
6. Les mêmes musiques que celles utilisées dans l'étude princeps étaient reprises dans notre étude,..... un extrait provenant de l'Allegro de Mozart (Petite Musique de nuit, Divertimento). [psy-art-246-body]	

- Remplissez les phrases suivante par **c'est-à-dire** ou **à savoir**.
- Est-ce que **c'est-à-dire** et **à savoir** sont interchangeables dans ces exemples ? Si non, expliquez quelle est la raison qui vous permet d'utiliser **c'est-à-dire** à la place de **à savoir**, et vice versa ? Vous pouvez vous appuyer sur les critères syntaxiques et sémantiques de chaque marqueur.
- Quels sont donc les emplois de chaque marqueur dans ces exemples ?

Approche : déductive + sémasiologique.

Explication :

À la suite des exercices sur les marqueurs *c'est-à-dire* et *à savoir*, l'enseignant fait travailler les apprenants sur un exercice lacunaire. Les apprenants ne se contentent pas de remplir les trous, il leur faut expliquer pourquoi ils décident d'utiliser *c'est-à-dire* ou *à savoir* dans chaque contexte. S'ils donnent de bonnes réponses, ils ont compris ce qui distingue ces deux marqueurs. Dans le cas contraire, l'enseignant peut intervenir et leur suggérer des indices sémantiques ou syntaxiques pour qu'ils trouvent la bonne réponse. À la fin, l'enseignant peut lister avec eux les caractéristiques principales et les spécificités de chaque marqueur. Il les aide à systématiser leurs connaissances.

Dans cet exercice, *c'est-à-dire* est utilisé pour introduire une reformulation à valeur métalinguistique, tandis que *à savoir* sert à énumérer ou à spécifier. De plus, au plan syntaxique, si *c'est-à-dire (que)* peut porter sur une proposition, ce n'est pas le cas de *à savoir*.

Nous allons maintenant décrire quelques fiches pédagogiques dont l'enseignant peut s'inspirer pour l'enseignement/apprentissage des MD à l'appui du corpus et également pour une initiation à l'écriture universitaire scientifique.

6.3.2.2. *Fiches pédagogiques*

En raison de la taille des articles scientifiques, nous avons décidé de ne pas faire travailler les apprenants sur des textes entiers, mais uniquement sur des passages sélectionnés pour leur pertinence. Précisons que les textes des trois premières fiches pédagogiques viennent d'un même article scientifique et que chaque extrait correspond à une partie textuelle stratégique : l'introduction, le développement et la conclusion. Par ces fiches pédagogiques, nous insistons essentiellement sur la mise en pratique de différentes approches pour développer les stratégies de lecture et de production chez les apprenants. Cela est d'autant plus important quand les étudiants vont travailler sur des textes des disciplines dont ils ne sont pas spécialistes.

L'apprentissage se déroule en plusieurs étapes. Ils peuvent se résumer en trois phases principales à savoir : 1/ découvrir les éléments pour prendre conscience de leur existence ; 2/ nommer et s'approprier ces éléments ; 3/ fixer et élargir le savoir. À chaque étape d'apprentissage, nous précisons le niveau des apprenants, les objectifs linguistiques et discursifs, et également l'approche ou la stratégie véhiculée, les différentes activités souhaitées. Nous allons également expliquer le rôle que joue l'enseignant à chaque étape. L'enseignant y trouvera des exemples pour l'introduction des concordanciers, la façon de les exploiter ainsi que des indications pour voir comment fonctionne un cours d'introduction au corpus. Nous aborderons également quelques exercices que l'enseignant peut introduire dans son cours.

Fiche pédagogique 1

Le texte est une introduction dans un article scientifique en psycholinguistique.

Dans de nombreuses langues, l'apprentissage et la mise en œuvre de l'orthographe grammaticale restent des problèmes majeurs, auxquels se trouvent confrontés les scripteurs enfants ou adultes. Dans le cas de l'orthographe du français, les erreurs sur l'accord en nombre des formes verbales écrites en /E/ (-ai(en)t, -é(es)) demeurent fréquentes au moins jusqu'à la classe de troisième (Brissaud et Chevrot, 2000). Plus généralement, des erreurs affectant la morphologie verbale subsistent même dans les écrits d'adultes de niveau de formation élevé (Lucci et Millet, 1994).

Les nombreuses étapes que doivent franchir les jeunes apprenants francophones ou anglophones pour résoudre les difficultés grammaticales de l'orthographe de leur propre langue commencent à être connues (Thevenin, Totereau, Fayol, 1997 ; Nunes, Bryant et Bindman, 1997). En revanche, on ignore presque tout de ce processus quand il s'agit de l'apprentissage de la morphologie d'une langue étrangère enseignée dans le cadre scolaire. L'acquisition du prétérit oral entre 2 et 6 ans et l'apprentissage de sa forme écrite ayant été largement décrits chez les jeunes anglophones, le choix de ce temps garantit la comparaison entre les processus d'apprentissage en langue maternelle et en langue étrangère.

En guise d'introduction, nous commencerons par présenter les principaux résultats concernant l'apprentissage du prétérit anglais par de jeunes Anglais, puis nous essayerons de cerner les conditions spécifiques caractérisant l'apprentissage de la morphologie verbale dans le cadre scolaire. Ensuite, nous proposerons une analyse contrastive des temps du passé en anglais et en français afin d'observer les secteurs de complexité potentielle pour les apprenants francophones. Enfin, nous présenterons une étude exploratoire de l'acquisition du prétérit anglais dans des textes écrits par des apprenants francophones. Nous la compléterons par une expérimentation construite à partir des hypothèses déduites de l'observation des formes verbales dans ces textes. Les résultats convergent pour suggérer que la conscience morphosyntaxique des apprenants, c'est-à-dire leur capacité à prendre conscience des unités morphosyntaxiques et à les manipuler délibérément, se centre sur la forme des verbes français au détriment des conditions d'usage du prétérit en anglais.

Coralie Payre-Ficout, Jean-Pierre Chevrot, La forme contre l'usage : étude exploratoire de l'acquisition du prétérit anglais. Lidil 30, 2001, 101-116

➤ **Niveau : B2/C1**

➤ **Objectifs :**

- Utiliser des marqueurs de topicalisation et d'enchaînement thématique
- Apprendre à rédiger une introduction
- Développer les stratégies de lecture

Afin d'aider l'enseignant à mieux traiter le texte, il peut se référer à la Figure 6-8, page 257. L'apprentissage se déroule en plusieurs étapes, l'enseignant peut articuler différentes approches en fonction du phénomène linguistique ou des objectifs visés.

➤ **1re étape : Observation**

Approche : inductive. Stratégie de lecture *top-down*.

Activités : Afin de sensibiliser les apprenants au rôle des MD dans le texte, l'enseignant enlève tous les MD du texte-source et leur demande dans un premier temps d'observer le texte sans MD. L'enseignant les fera réfléchir pour lister avec eux ce qu'il manque et vérifier s'ils connaissent les MD. S'ils en proposent, cela pourrait les aider à mieux comprendre de quoi il s'agit. Ensuite, il leur donne le texte originel avec les MD, les apprenants peuvent comparer les deux documents, avec et sans les MD. La signification des MD monolexicaux *puis*, *ensuite*, *enfin* à ce niveau de langue française (B2/C1) est déjà acquise par le public visé.

Explication : Cette étape correspond en fait à une première lecture du texte. L'activité de compréhension écrite et de sensibilisation est selon nous très importante et peut s'appliquer à tous les niveaux. Il est important de commencer par une telle activité pour que l'apprenant prenne conscience lui-même de l'importance que jouent les MD. Sans les MD, le texte manque de cohérence et n'est qu'une juxtaposition d'idées, et le lecteur risque de perdre le fil tout au long du texte.

À la suite de cette étape, l'apprenant va analyser le texte sur deux plans : le plan syntaxique grâce aux MD et le plan sémantique grâce aux structures discursives.

➤ **2e étape : Repérage des marqueurs de topicalisation (aspect syntaxique)**

Approche : Approche inductive, aspect syntactico-sémantique.

Activités : L'enseignant prépare un exercice tiré d'un concordancier de *en guise de*, et propose à l'apprenant d'observer dans un premier temps l'aspect syntaxique de ce MD. Il peut s'inspirer du concordancier ci-dessous.

- *Quelle est la structure syntaxique de ces MD ?*
- *Quel est l'élément qui précède et qui suit le MD ?*
- *Quelle est la position du MD ?*

Extrait du concordancier de *en guise de* :

En guise d' introduction, nous commencerons par présenter les principaux résultats concernant l'apprentissage du prétérit anglais par de jeunes Anglais, puis nous essayerons de cerner les conditions spécifiques caractérisant l'apprentissage de la morphologie verbale dans le cadre scolaire. [lin-art-225-introduction]
En guise d' illustration, je présente une étude en cours sur l'énumération, qui exemplifie les prises de position exposées jusque là. [lin-hdr -364-body]
J'esquisse en guise de conclusion quelques grandes lignes de ces évolutions, de façon partielle et inévitablement partielle, en lien avec mes recherches. [lin-hdr -364-conclusion]
Quel que soit le rôle et la plus ou moins grande implication de ces collaborateurs, nous tenons à les citer en guise de remerciement : Nadine Boissié, Jérôme Faux, Thierry Lamarque, Nathalie Morales, Françoise Roland, Maïté Rouire, Viviane Trignac. [lin-com-519-notes]
Signes en général (Section 2.3), nous allons rappeler brièvement le traitement du signe chez Saussure (Section 2.1) et nous attarder sur le traitement des signes proposé par Igor Mel'auk (Section 2.2), qui préfigure celui de GUST. La Section 2.4, en guise de conclusion, statuera sur la question de la double articulation de la langue. [lin-hdr -615-body]

Explication : L'apprenant se trouve toujours dans une phase de découverte. Cette activité permet à l'apprenant de mieux repérer la structure syntaxique des marqueurs de topicalisation, et éventuellement leur position (en tête de phrase). En raison de sa structure syntaxique, *en guise de* peut s'associer avec un nom. Pourtant, en observant les exemples, l'apprenant peut constater que *en guise de* n'accepte que quelques noms qui renvoient à une partie textuelle spécifique (introduction, conclusion, remerciement). L'enseignant peut développer les stratégies de lecture *bottom-up*.

➤ 3^e étape : Repérage des marqueurs de topicalisation (aspect sémantique)

Approche : inductive. Stratégie de lecture : *top-down*.

Activités : Il fait observer le rapport entre le contexte de gauche et le contexte de droite du MD pour découvrir le rôle des marqueurs de topicalisation. L'enseignant peut s'inspirer des concordances des marqueurs de topicalisation dans Dicorpus pour créer des exercices. Il s'agit d'exercices lacunaires (comme dans l'exercice 3 sur *c'est-à-dire* et *à savoir* présenté ci-dessus), d'exercices d'appariement, etc.

L'enseignant fait observer le contexte gauche et droit du marqueur et demande de relever les différences d'emploi de *en guise de*.

Explication : Cette étape permet de comprendre quel est le rôle exact des marqueurs de topicalisation et comment ils participent à la structuration du texte. Ces marqueurs servent à

introduire une nouvelle thématique. Toutes ces activités servent à donner à l'apprenant des moyens lexicaux pour passer à un nouveau contenu ou une nouvelle idée.

➤ **4^e étape : Classification des MD**

Approche : sémasiologique. Stratégie de lecture : *bottom-up*.

Activités : L'enseignant peut relever avec l'apprenant tous les MD du texte et les classer en fonction de leur sens. Il s'agit des marqueurs servant à présenter le plan de développement : *puis, ensuite, enfin*, les marqueurs de topicalisation : *en guise de*, le marqueur de reformulation : *c'est-à-dire*. L'enseignant peut faire un exercice sous forme de question – réponse.

Explication : L'apprenant entre dans la deuxième phase d'appropriation des éléments linguistiques. Après des activités de découverte des propriétés syntaxiques et sémantiques des MD, cette activité leur permet de systématiser le rôle de chaque groupe de MD.

➤ **5^e étape : Repérage des structures rhétoriques**

Approche : déductive + inductive. Stratégie de lecture : *top-down*.

Activités : Pour faciliter la compréhension, l'enseignant peut expliquer à l'apprenant quelques termes difficiles : *prétérit, morphologie, morphosyntaxique*. Pour guider l'apprenant dans l'analyse des structures discursives, l'enseignant prépare quelques questions : *Quel est le sujet de recherche ? Quel est l'intérêt de cette recherche ? Comment les auteurs envisagent-ils de développer les idées dans le texte ?* L'apprenant peut s'appuyer sur les MD pour relever les idées essentielles dans l'extrait.

Ensuite, l'enseignant demande aux apprenants de résumer/rassembler les idées principales à aborder dans l'introduction, à savoir le contexte de recherche, le thème de recherche, l'implication personnelle, la problématique, le choix méthodologique et l'annonce du plan.

Explication : L'observation des MD effectuée à l'étape précédente permet à l'apprenant de comprendre la valeur sémantique des MD, ce qui facilite en fait le repérage des idées essentielles et des structures discursives développées dans le texte. La structure de cette introduction est assez claire. Les auteurs commencent par expliquer le contexte de recherche,

les raisons qui motivent leur recherche. L'utilisation du marqueur de topicalisation *en guise de* marque un changement de thématique où les auteurs présentent le plan de travail. L'apprenant repère par la suite une série des marqueurs monolexicaux qui signalent différentes étapes que les auteurs envisagent de développer. L'apprentissage se déroule selon une approche onomasiologique.

➤ **6^e étape : Production écrite**

Aspect : rhétorique, lexical et sémantique.

Activités : L'enseignant demande à l'apprenant de rédiger une introduction sur son propre sujet de recherche en lui donnant une suite des questions auxquelles l'apprenant doit répondre pour construire son texte. L'enseignant peut reprendre les mêmes questions de l'étape précédente en lui demandant d'utiliser obligatoirement un MD. L'apprenant doit réutiliser les marqueurs de topicalisation, et également les marqueurs d'enchaînement thématique qu'il aura travaillés dans cette leçon.

- *Quel est le sujet de recherche ? Utilisez un marqueur servant à présenter le plan de développement.*
- *Quel est l'intérêt de cette recherche ? Vous devez utiliser un marqueur de topicalisation.*
- *Comment les auteurs envisagent-ils de développer les idées dans le texte ? Utilisez un marqueur de reformulation.*
- *Remettez ces idées dans le bon ordre.*

Explication : La production écrite joue un rôle important dans la consolidation des connaissances. Par les activités de compréhension écrite, l'apprenant observe, analyse des MD et repère des structures discursives dans le texte. Il prend conscience de la façon dont les auteurs présentent leurs idées dans le texte, et par les activités de production écrite, l'apprenant peut réinvestir ses acquis. L'apprenant peut utiliser le Dicorpus pour l'aider dans la phase de rédaction.

À l'étape précédente, l'apprenant a appris à découper le texte. À cette étape, il doit en construire un à leur manière au travers d'un sujet personnel, mais en réutilisant les MD qu'ils ont vu et appris. Toutes les consignes servent à guider l'apprenant en rédigeant son texte.

Consigne :

Dans la Figure 6-8 : Schéma d'analyse - Fiche pédagogique 1, nous récapitulons les idées essentielles du texte. Les parties discursives sont encadrées en bleu. Ce schéma clarifie les trois parties distinctes qui abordent une idée essentielle dans l'introduction. Tous les MD sont également soulignés : les MD à valeur métadiscursive sont mis en vert et les MD à valeur argumentative en jaune.

À l'étape précédente, l'apprenant a appris à découper le texte. À cette étape, il doit en construire un à leur manière en fonction du sujet personnel mais en réutilisant les MD qu'ils ont vus et appris dans le texte. Toutes les consignes servent à guider l'apprenant en rédigeant son texte.

Consigne :

Dans la Figure 6-8 : Schéma d'analyse - Fiche pédagogique 1, nous essayons de schématiser les idées essentielles du texte. En fait, les parties discursives sont encadrées en bleu. Ce schéma clarifie les trois parties distinctes dans cette introduction, chaque partie traite une idée essentielle à aborder dans l'introduction. Tous les MD sont également soulignés : les MD à valeur métadiscursive sont mis en vert et les MD à valeur argumentative en jaune.

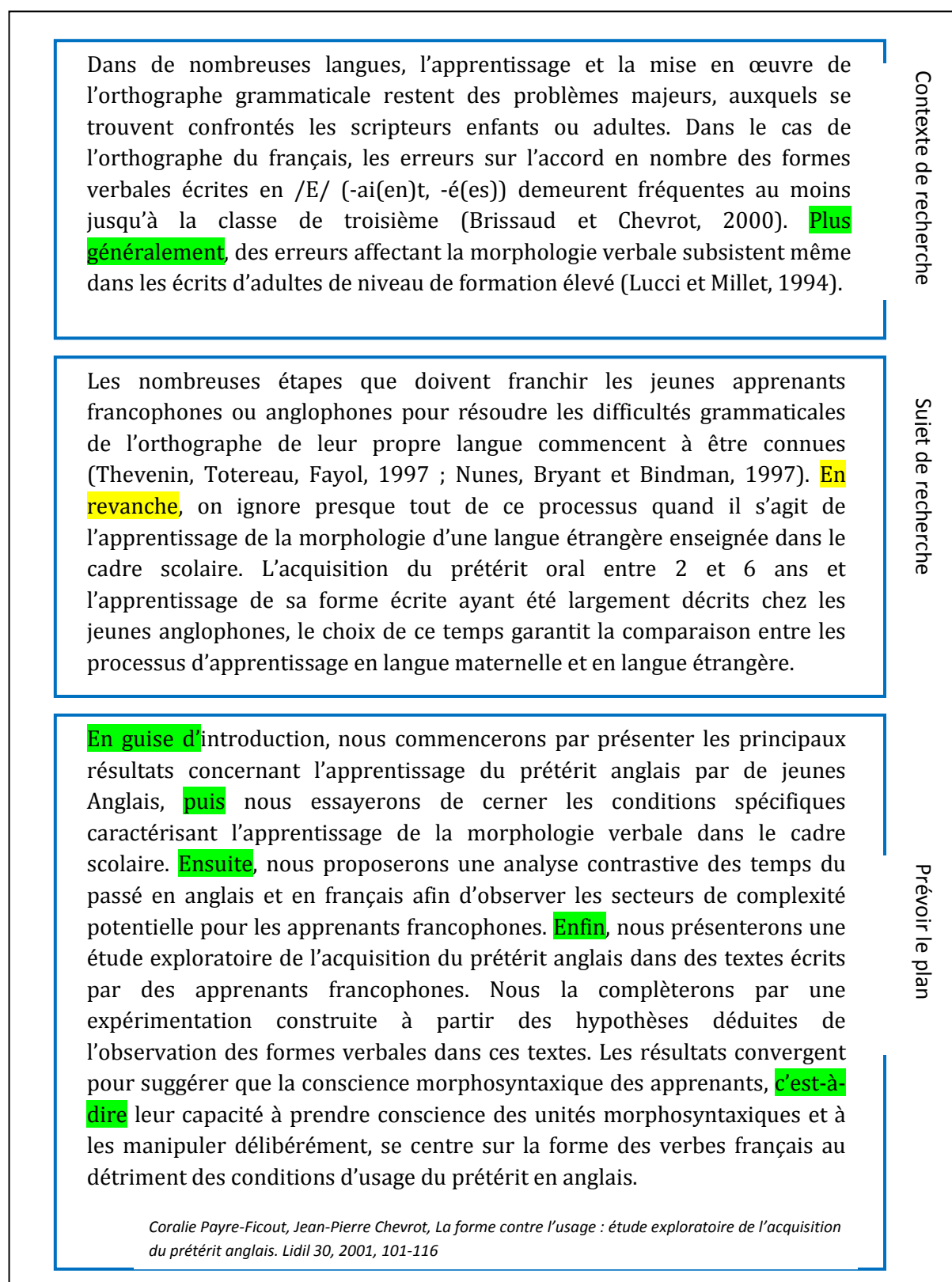


Figure 6-8 : Schéma d'analyse - Fiche pédagogique 1

Fiche pédagogique 2

Cet extrait se trouve dans le développement de l'article en psycholinguistique.

S'agissant des temps du passé, l'anglais et le français sont différents au niveau de l'organisation des paradigmes et de leur usage.

Les deux systèmes d'aspect et de temps ne sont pas superposables. Si l'on considère par exemple la notion de passé, on est loin d'une correspondance terme à terme entre les paradigmes des deux langues. Selon le contexte, au prétérit anglais peut correspondre l'usage du passé composé, de l'imparfait et du passé simple. Au present perfect, correspondent le passé composé et le présent. Si l'on part du français, comme le font sans doute les apprenants débutants de l'anglais, les correspondances sont aussi complexes. En face du passé composé, il faudra choisir entre un prétérit et un present perfect. Seul le passé simple renvoie de façon univoque au prétérit.

Sur le plan de la forme, la langue française est une langue flexionnelle riche, notamment à l'écrit, caractérisée par un décalage entre la morphologie de l'oral et celle de l'écrit (Jaffré, 1998). Plus précisément, l'écrit distingue des formes verbales qui sont homophones à l'oral. La connaissance de la morphologie verbale orale ne permet donc pas de prédire les formes écrites, la transcription phonographique directe de l'oral ne prenant en compte ni l'existence de morphogrammes muets ni la multiplicité des liens entre phonèmes et graphèmes (dans le cas des formes verbales homophones en /E/ par exemple). Ainsi, la grammaire joue un rôle primordial dans la sélection de la flexion écrite. A l'inverse du français, l'anglais présente un système flexionnel extrêmement pauvre. La langue anglaise ne compte que deux temps grammaticaux, le présent et le passé, et ne connaît, à l'exception de Be et de quelques verbes irréguliers, que trois marqueurs de la forme simple, deux pour le présent simple (-set Ø) et un seul pour le prétérit (-ed). On peut trouver un verbe anglais sous quatre formes s'il est régulier (play, plays, played, playing) et entre trois et cinq s'il est irrégulier (Lapaire et Rotgé, 1991 ; Larreya et Rivière, 1991). A l'inverse du français, la graphie des verbes anglais neutralise souvent des différences sonores (cf. les différentes prononciations de la terminaison -ed). De ce fait, à la différence du français, toute flexion écrite de l'anglais a un correspondant oral, ce qui facilite la scription. Toutefois, à une flexion écrite unique peuvent correspondre différentes prononciations. Comme en français, la transcription phonographique ne garantit donc pas la justesse morphographique.

En définitive, dans les deux langues, les apprenants doivent apprendre à se détacher de la forme orale pour produire la forme écrite, mais ils procèdent de manière différente. Les apprenants de la morphographie du français doivent plutôt comprendre qu'une même séquence sonore peut s'écrire de façons différentes. A l'inverse, les apprenants de la morphographie de l'anglais doivent plutôt comprendre que des séquences sonores différentes peuvent s'écrire de la même façon.

Ces différences dans l'usage et la forme des temps passés des deux langues devraient se manifester par des erreurs concernant ces deux plans dans les textes des jeunes Français apprenant l'anglais.

Coralie Payre-Ficout, Jean-Pierre Chevrot, La forme contre l'usage : étude exploratoire de l'acquisition du prétérit anglais. Ldil 30, 2001, 101-116

➤ **Objectifs linguistiques :**

- Développer une idée
- Marqueurs de topicalisation, marqueurs de comparaison, marqueurs de reformulation

La Figure 6-9 : Schéma d'analyse - Fiche pédagogique 2 représente dans un schéma les idées principales du texte. L'enseignant peut l'utiliser pour créer de multiples exercices. Nous proposons plusieurs étapes pour la réalisation de cette activité.

➤ **1^{re} étape : Observation**

Approche : inductive, aspect lexical.

Activités : L'enseignant demande aux apprenants d'observer le texte sans MD et les laisse réfléchir à ce qu'il manque. Il peut lister avec eux leurs suggestions. S'ils proposent des MD, il faudra les guider, si nécessaire, pour les classer et vérifier que les choix de MD sont pertinents au plan sémantique et syntaxique.

L'enseignant peut alors demander aux apprenants de comparer leur texte complété avec le texte original. L'apprenant observe la différence entre les deux textes au niveau de l'organisation textuelle et une discussion peut suivre sur les choix opérés par les auteurs et par les apprenants.

Explication : Cette activité s'inscrit dans une phase de découverte. Elle est très bénéfique à la mémorisation quasi-immédiate. De plus, l'activité de comparaison permet de sensibiliser l'apprenant à la fonction d'organisation qu'apportent les MD.

➤ **2^e étape : Marqueurs de comparaison**

Approche : déductive.

Activités : L'enseignant fait travailler l'apprenant sur les marqueurs de comparaison : *à l'exception de*, *à la différence de*, *comme*. L'enseignant demande de relever la structure syntaxique de ces marqueurs et également la nature syntaxique.

Explication : L'apprenant part de quelque chose qui peut paraître plus simple à décoder. L'enseignant peut éventuellement se passer de cette étape si l'apprenant peut déjà relever ces marqueurs à la première étape.

➤ **3^e étape : Repérage du marqueur de reformulation (aspect syntaxique) : *en définitive***

Approche : inductive, aspect syntaxique.

Activités : L'enseignant extrait des exemples du concordancier de *en définitive*. Il demande à l'apprenant de relever les propriétés syntaxiques du marqueur. L'enseignant peut lui faire remplir un tableau (comme dans Exercice 1 et 2 sur *c'est-à-dire* et *à savoir*) pour que l'apprenant puisse mieux voir les propriétés syntaxiques de ce marqueur.

Or les motifs composant la classe dévoiement de la fonction policière ont tous en commun l'emploi, le signalement, la mise en évidence ou le recours indus aux objets ou aux pouvoirs, en définitive, aux « choses » qui incarnent concrètement le métier policier. [spo-art-380-body]
Finalement, les jeunes chercheurs, par autocensure, laissent souvent les seuls chercheurs avérés et reconnus (c'est-à-dire bien souvent des chercheurs seniors) utiliser et, en définitive, s'approprier le forum de discussion mis en place par la revue, dans un phénomène de reproduction d'une hiérarchie institutionnelle déjà bien installée. [infcom-art-432-body]
Nous verrons que ces trois modalités de partage sous- tendent en définitive une tension entre les dimensions individuelle et collective de la consommation sérielle, que cette population de fans fait apparaître de manière saillante. [infcom-art-21-body]
Mais on va le voir, c'est également un moyen de reprendre le contrôle de la production, de redéfinir le rôle qu'ils y jouent et, en définitive, de tenter de revaloriser l'image qu'ils ont d'eux-mêmes celle qu'ils renvoient. [soc-art-64-body]
En effet, dans la perspective du projet HarmoS, les standards doivent être décrits en termes de performances par rapport à des tâches et/ou situations ; ils correspondent au produit de la compétence dans une situation donnée ; leur établissement passe par conséquent par la définition de tâches se référant au modèle de compétence et l'observation des performances des élèves qui permettra en définitive de vérifier la validité du modèle même. [sed-com -26-body]

Explication : Cette étape a pour rôle d'orienter l'apprenant pour se concentrer dans un premier temps sur les aspects syntaxiques de *en définitive*. L'enseignant peut se servir du Tableau 5.12 pour avoir les propriétés essentielles de ce marqueur.

➤ **4^e étape : Repérage du marqueur de reformulation (aspect syntaxique) : *en définitive***

Approche : inductive, aspect sémantique.

Activité : L'enseignant demande à l'apprenant d'analyser la relation entre le contexte de gauche et le contexte de droite de *en définitive*. Cette activité permet à l'apprenant de prendre conscience des valeurs sémantiques de ce marqueur.

Cette activité peut être suivie d'un exercice de révision qui est susceptible d'aider l'apprenant à consolider ses connaissances.

Explication : *En définitive* sert à mettre en lien une série d'idées et le point de vue de l'auteur est considéré comme définitif. Le fait de travailler sur un seul marqueur aide l'apprenant à mieux comprendre son fonctionnement.

➤ 5^e étape : Marqueurs de reformulation *en définitive/en fin de compte*

Approche : inductive + déductive, aspect syntactico-sémantique.

Activités : L'enseignant met en comparaison les deux marqueurs de reformulation *en définitive* et *en fin de compte*. Il peut se servir du tableau de comparaison des caractéristiques des marqueurs de reconsidération dans le chapitre 5 (Tableau 5-15). L'enseignant systématise ou explicite ce qui distingue ces deux marqueurs. Il peut également proposer d'autres exercices, des exercices lacunaires ou des exercices de production écrite où l'apprenant doit reformuler deux ou plusieurs énoncés tout en utilisant *en définitive* ou *en fin de compte*.

Nous proposons ici d'introduire un exercice lacunaire tout en demandant aux apprenants d'expliquer leur choix de MD.

<i>En définitive/en fin de compte</i>	Explication
1. Hjelmslev (1966) écrit : « une typologie linguistique exhaustive est la tâche la plus grande et la plus importante qui s'offre à la linguistique., sa tâche est de répondre à la question : quelles structures linguistiques sont possibles ? et pourquoi telles structures sont possibles quand d'autres ne le sont pas ? [lin-the -147-body]	
2. Finalement, les jeunes chercheurs, par autocensure, laissent souvent les seuls chercheurs avérés et reconnus (c'est-à-dire bien souvent des chercheurs seniors) utiliser et,....., s'approprier le forum de discussion mis en place par la revue, dans un phénomène de reproduction d'une hiérarchie institutionnelle déjà bien installée. [infcom-art-432-body]	
3. Plus un appellatif est général, plus l'énoncé doit contenir d'informations qui restreignent son extension et,....., justifie son emploi. [lin-com-518-body]	
4. La faculté de langage devrait être ancrée aussi bien dans les perceptions (base des catégories, comme dans de nombreux modèles) que dans les valeurs (catégorisation fondée sur la valeur, et..... fonction du langage) et dans l'action	

(ordonnancement, performatifs). [tal-the -641-body]	
5. Il en résulterait une nouvelle planification de l'activité, de nouvelles priorités et,..... une redistribution des ressources disponibles. [psy-art-407-body]	
6. Autre exemple, ses nombreux travaux sur le pronom <i>il</i> , qui semblent défendre..... un mode de fonctionnement pragmatique, du moins dans certains cas ; nous présenterons cependant ci-dessous un modèle plus achevé de traitement pragmatique de la référence (V.D. 5). [tal-the -641-body]	

Réponses :

1. En fin de compte, 2. En définitive, 3. En fin de compte, 4. En fin de compte, 5. En définitive, 6. En fin de compte

Explication : Il s'agit d'une activité facultative. En fonction du niveau de l'apprenant et du programme d'apprentissage, l'enseignant peut utiliser l'activité de comparaison dans cette leçon, si l'apprenant connaît déjà les propriétés syntaxiques et sémantiques de *en fin de compte*. Si non, il peut intégrer cette étape à une autre leçon qui se consacre essentiellement à l'analyse de *en fin de compte* par exemple. L'enseignant recourt à cette activité de comparaison pour faire réviser l'apprenant sur les propriétés syntaxiques et sémantiques de deux marqueurs qui sont souvent considérés comme synonymes. Cette activité a pour rôle de consolider les connaissances de l'apprenant et de l'aider à distinguer *en définitive* de *en fin de compte*. Il est nécessaire d'expliquer clairement les propriétés syntaxiques et sémantiques qui différencient ces deux marqueurs, ce qui manque encore dans les méthodes de FLE. En effet, *en définitive* souligne un point de vue considéré généralement comme définitif sur un sujet, tandis que *en fin de compte* introduit un point de vue de valeur terminative sur une question traitée, mais pas sur le sujet entier. Le premier est « global » au plan conceptuel et l'autre « local » sur une question à un moment donné.

➤ 6^e étape : Classement des MD

Approche : déductive + sémasiologique. Stratégies de lecture *bottom-up*

Activités : L'enseignant demande à l'apprenant de repérer tous les MD dans cet extrait et de les regrouper selon leur fonction rhétorique dans l'énoncé. On peut relever les marqueurs de comparaison : *à l'exception de*, *à la différence de*, *à l'inverse de*, *comme*, etc. ; les marqueurs de topicalisation : *sur le plan de*, *en guise de* ; les marqueurs de reformulation : *plus précisément*, *en définitive*, les marqueurs de concession : *toutefois*, *mais*, l'exemple d'exemplification : *par exemple*, le marqueur de causalité : *de ce fait*, *ainsi*.

Explication : Cet extrait utilise essentiellement des marqueurs de comparaison et d'exemplification, car les auteurs souhaitent expliquer l'usage des temps du passé entre deux systèmes de langue : le français et l'anglais. Cette activité de repérage permet une prise de conscience de ces marqueurs notamment par la fréquence élevée des marqueurs de comparaison dans cet exemple. L'apprenant peut prévoir dans une certaine mesure les idées que vont développer les auteurs ; ceci pourrait d'ailleurs être l'occasion d'une activité collective.

➤ **7^e étape : Consolidation des connaissances sur les MD**

Approche : déductive

Activités : L'enseignant peut créer différents exercices sur les marqueurs de comparaison et de topicalisation : exercices lacunaires, exercices d'appariement en se basant sur les concordances dans Dicorpus.

Par exemple, les marqueurs de topicalisation (*en guise de*) et les marqueurs de comparaison (*à l'instar de*, *à la différence de*, *à l'exception de*) portent sur un groupe nominal. À la suite des exercices dans la Fiche pédagogique 1, les apprenants peuvent conclure que les marqueurs de topicalisation comme *en guise de* ne portent que sur quelques noms spécifiques. Nous suggérons de proposer des exercices lacunaires pour une aide à la mémorisation (ils servent notamment de rappel) tout en leur demandant des explications du choix du MD (comme l'exercice à la 5^e étape sur *en définitive/en fin de compte*).

Explication : Cette étape permet de consolider les connaissances des apprenants sur ces MD. L'apprenant devrait alors savoir compléter les phrases tout en donnant de bonnes explications et l'enseignant pourra en déduire qu'il a mémorisé le fonctionnement du MD.

➤ **8^e étape : Repérage des idées essentielles**

Approche : inductive + onomasiologique. Stratégie de lecture *top-down*

Activités : L'apprenant doit relever les idées essentielles dans cet extrait grâce aux MD repérés à l'étape précédente.

- *Comment les idées sont-elles organisées pour progresser ?*
- *Quelles sont les principales différences entre les temps du passé en anglais et en français ?*

- *Quelle est la conclusion des auteurs concernant ces différences ?*

Une fois les idées relevées, l'apprenant les classe dans l'ordre du texte. L'enseignant suggère à l'apprenant de repérer comment elles sont introduites linguistiquement.

Explication : À cette étape, l'apprenant commence à traiter le texte au plan sémantique. Chaque MD est relatif à une fonction d'écriture spécifique, par exemple : reformuler pour les marqueurs de reformulation, comparer pour les marqueurs de comparaison, introduire un nouveau contenu pour les marqueurs de topicalisation, etc. La prise en compte préalable du rôle de ces MD permet à l'apprenant de repérer plus facilement les idées essentielles dans le texte.

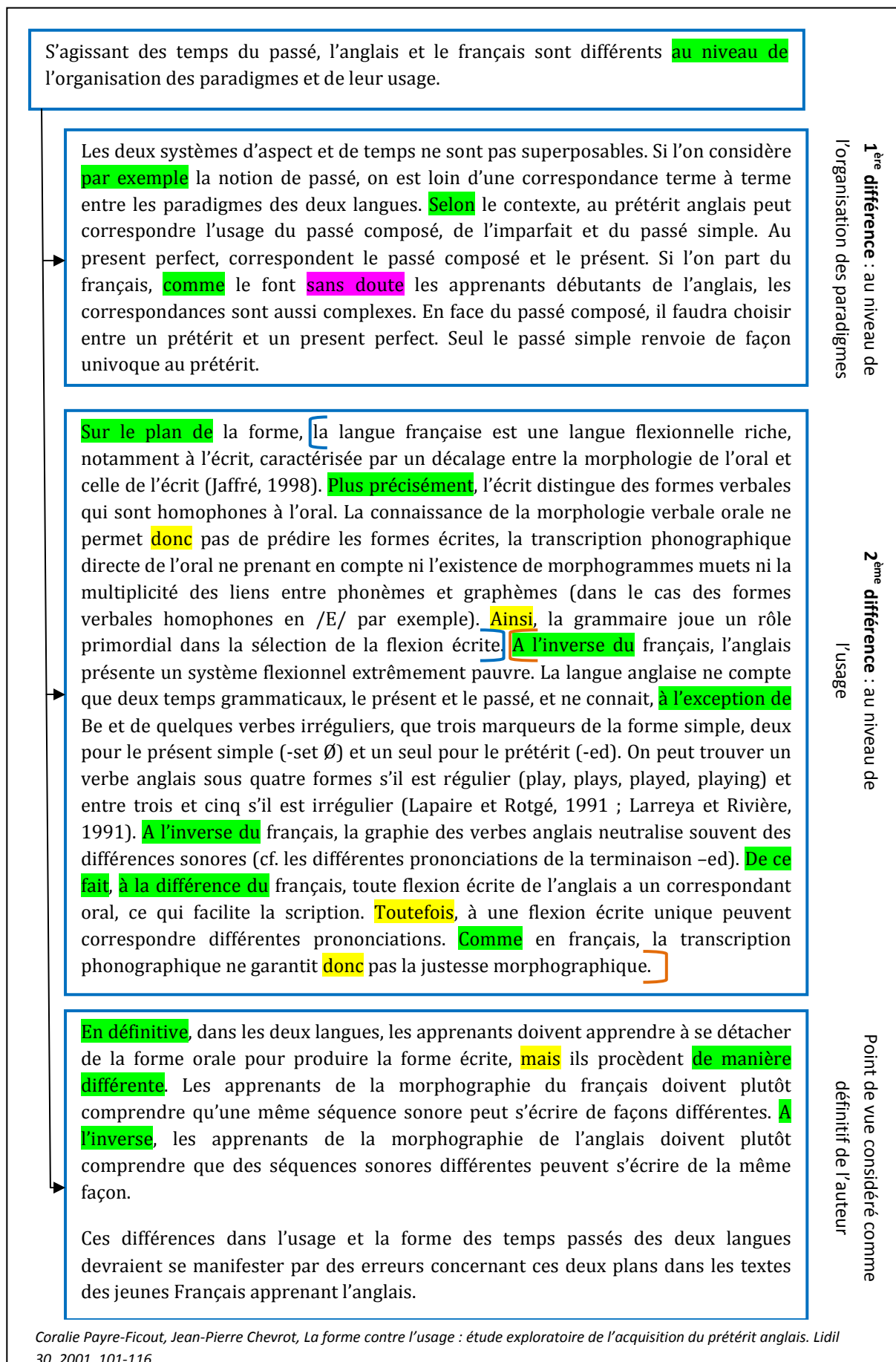
➤ **9^e étape : Production écrite**

Aspect : rhétorique, syntactico-sémantique.

Activités : L'apprenant doit rédiger un texte d'une quinzaine de lignes dans le domaine de son choix. Il doit utiliser les marqueurs abordés dans la leçon : marqueurs de comparaison, de topicalisation et de reformulation.

Explication : Grâce aux exercices de production écrite, l'apprenant peut exploiter à nouveau les connaissances qu'il aura acquises y compris les moyens lexicaux, les fonctions rhétoriques des écrits scientifiques, etc.

Dans la Figure 6-9 : Schéma d'analyse - Fiche pédagogique 2, nous essayons de représenter de manière schématique les idées essentielles développées par l'auteur. Cette représentation permet de montrer plus clairement à l'enseignant et à l'apprenant le rôle des MD dans l'organisation des idées et que l'enseignement/apprentissage des MD sert à repérer les idées essentielles du texte. Cet extrait d'article est bien structuré. La première phrase annonce deux différences sur *les temps du passé* en anglais et en français, les deux paragraphes qui suivent servent à expliciter ces différences. Le marqueur de topicalisation *Sur le plan de* souligne un changement de thématique. On peut distinguer les MD grâce à différentes couleurs : le vert pour les métadiscursifs, le jaune pour les argumentatifs, les séquences lexicalisées à fonction modale sont indiquées en violet. Les crochets servent à marquer l'étendue de chaque idée.



Coralie Payre-Ficout, Jean-Pierre Chevrot, *La forme contre l'usage : étude exploratoire de l'acquisition du prétérit anglais*. Lidil 30, 2001, 101-116

Figure 6-9 : Schéma d'analyse - Fiche pédagogique 2

Fiche pédagogique 3

Cet extrait correspond à la partie « discussion » d'un article scientifique en psycholinguistique.

Nous constatons que l'acquisition du prétérit anglais par des apprenants français diffère de son acquisition par les jeunes Anglais. En effet, certains des résultats obtenus par l'analyse du corpus et par le protocole expérimental ne correspondent pas aux points forts du scénario développemental élaboré par Terezinha Nunes et ses collègues (1997).

Premièrement, les erreurs de régularisation, fréquentes chez les jeunes anglophones, sont quasi inexistantes dans les textes des jeunes Français mais nombreuses dans l'expérimentation. La confrontation entre un recueil contrôlé des données par traduction et un recueil plus spontané par narration écrite indique que la régularisation est un processus disponible chez les apprenants français, mais peu exploité en production libre de textes, sans doute du fait de l'évitement des verbes irréguliers peu ou pas connus.

Deuxièmement, nous observons un phénomène massif, dont l'origine n'a jamais été établie empiriquement dans les études sur l'apprentissage du prétérit anglais par les jeunes Anglais : le remplacement d'une forme simple par une forme composée. L'analyse du corpus indique également que l'auxiliaire impliqué dans les erreurs que nous avons nommées «be conjugué» et «have conjugué» correspond dans 69 % des cas à l'auxiliaire français.

Nous interpréterons ce classique phénomène de transfert dans les termes d'une focalisation de la conscience morphosyntaxique des sujets français sur la forme au détriment de l'usage. Les apprenants chercheraient à reproduire la séquence «auxiliaire + participe passé» du français et négligeraient les conditions d'usages des temps de l'anglais. On peut supposer que cette centration sur la forme au détriment de l'usage provient de l'absence d'un enseignement de l'aspect dans la grammaire scolaire du français. Plus précisément, la distinction entre le passé composé à valeur d'accompli du présent, qui se traduit par un present perfect anglais, et le passé composé à valeur aoristique, qui se traduit par un prétérit anglais, ne fait pas l'objet d'un enseignement systématique à l'école primaire et au collège. Démunis des bases métaagrammaticales leur permettant de comprendre les conditions d'usage des temps de l'anglais, les jeunes Français utiliseraient la forme composées ou la forme simple du français comme unique repère. Ce serait donc bien un défaut de la conscience morphosyntaxique dans la langue maternelle qui provoquerait l'erreur dans la langue étrangère.

Dans le domaine de la langue maternelle, on connaît depuis longtemps l'impact de la conscience métalinguistique sur les apprentissages scolaires de l'écrit (Gombert, 1990). Plusieurs études ont démontré également que les connaissances métalinguistiques en langue maternelle sont étroitement liées à la réussite de l'apprentissage de la langue étrangère. Ce résultat est établi pour l'apprentissage de l'hébreu par des enfants anglais et l'apprentissage de l'anglais par des enfants portugais, même si les langues en jeu ont des structures différentes (Bindman, 2003 ; Castro, 2003). Ce transfert ne s'effectuerait pas seulement de la langue maternelle vers la langue étrangère. En effet, les principes grammaticaux acquis en langue étrangère peuvent apporter une meilleure compréhension aux enfants qui sont en train de développer leur conscience grammaticale en langue maternelle. Ceci peut s'avérer vrai pour des aspects grammaticaux qui sont marqués plus clairement en langue étrangère qu'en langue maternelle et qui bénéficient d'une instruction explicite.

Dès lors, nous rejoignons les résultats précurseurs et les positions défendues par Louise Dabène (1983) dans un cadre didactique : un travail métalinguistique sur la grammaire de la langue maternelle en situation d'apprentissage scolaire doit contribuer à une meilleure maîtrise du système de la langue étrangère.

Coralie Payre-Ficout, Jean-Pierre Chevrot, La forme contre l'usage : étude exploratoire de l'acquisition du prétérit anglais. Lidil 30, 2001, 101-116

➤ **Niveau : B2/C1**➤ **Objectifs :**

- Marqueurs de récapitulation, moyens pour exprimer une hypothèse
- Rédiger une conclusion

Nous essayons de mettre en relief les rapports entre les parties dans le texte ainsi que les idées essentielles dans la Figure 6-10 : Schéma d'analyse - Fiche pédagogique 3.

➤ **1^{re} étape : Repérage des MD (aspects syntaxique et sémantique)**

Approche : déductive + sémasiologique. Stratégie de lecture *bottom-up*.

Activités : Dans un premier temps, l'enseignant demande à l'apprenant de relever les MD dans le texte et voir avec lui ceux qu'ils connaissent déjà. Dans un deuxième temps, il lui demande de les regrouper selon leur fonction rhétorique : marqueur d'enchaînement thématique (*premièrement, deuxièmement*), marqueur de reformulation (*dès lors*), marqueurs de causalité (*du fait de, en effet*), marqueurs de valeur conclusive (*dès lors, donc*), et les marqueurs exprimant l'hypothèse (*on peut supposer que*).

Explication : Cette étape consiste à aborder le texte au plan structurel. La prise de conscience des valeurs sémantiques des MD sert à segmenter la phrase par différents « constituants » et idées, ce qui facilite la compréhension du texte et permet une lecture plus active.

➤ **2^e étape : Repérage les fonctions rhétoriques**

Approche : inductive. Stratégie *top-down*

Activités : L'enseignant pose des questions pour que l'apprenant révèle les idées essentielles de chaque texte.

- *Quels sont les résultats présentés dans le texte ?*
- *Existe-t-il des différences entre les résultats de cette recherche et les travaux précédents ?*
- *Quels sont les rôles de la conscience métalinguistique dans l'acquisition d'une langue étrangère ?*
- *Quelle est la conclusion finale des auteurs ?*

L'enseignant guide l'apprenant dans le repérage des idées essentielles en lui suggérant de se baser sur les MD.

- *Comment les MD servent-ils à repérer les idées essentielles dans la conclusion ?*
- *Combien d'idées y a-t-il dans cet extrait ?*

Il est important que l'enseignant précise à cette étape les idées attendues dans une conclusion : présenter des résultats, des limites et des perspectives.

Explication : Grâce aux MD, l'enseignant demande à l'apprenant de repérer les « zones argumentatives ». L'apprenant peut recourir à d'autres connaissances sur la langue (anaphore) pour repérer les idées essentielles. Dans la Figure 6-10 : Schéma d'analyse - Fiche pédagogique 3, nous montrons d'une manière schématique comment les idées sont reliées entre elles dans le texte. Les anaphores sont mises en rouge et soulignées, il s'agit d'un des phénomènes qui guidera l'apprenant dans le repérage des idées.

➤ **3^e étape : Moyens pour introduire une hypothèse**

Approche : onomasiologique + inductive

Activités : L'enseignant s'appuie sur les concordances des marqueurs qui servent à introduire une hypothèse dans Dicorpus pour créer des exercices : exercices lacunaires, exercices de reformulation ou de création, etc. L'enseignant enlève les marqueurs qui introduisent une hypothèse, et demande à l'apprenant de relier les deux énoncés pour exprimer une hypothèse. L'apprenant peut accéder à d'autres exemples de l'utilisation de ces MD dans Dicorpus s'il le souhaite.

En outre, sur le site du corpus Scientext, il existe des grammaires préétablies par les membres de l'équipe Lidilem sur les collocations exprimant l'hypothèse. L'enseignant peut s'appuyer sur ces concordanciers pour créer différents exercices.

Explication : Cette étape se trouve à la deuxième phase d'apprentissage. Après des étapes de repérage, cette activité permet à l'apprenant de s'approprier des moyens lexicaux qui expriment une hypothèse.

➤ **4^e étape : Marqueur de récapitulation dès lors**

Approche : inductive

Activités : L'enseignant fait observer à l'apprenant quelques occurrences de *dès lors* et lui demande de relever le fonctionnement syntaxique et sémantique de ce marqueur.

Explication : En raison de sa structure syntaxique, ce marqueur peut poser quelques difficultés à l'apprenant. Il est important que celui-ci prenne conscience de ces différentes valeurs syntaxiques.

➤ **5^e étape : Production écrite**

Approche : rhétorique

Activités : L'enseignant demande à l'apprenant de rédiger une conclusion sur le sujet de son choix tout en prenant en considération des contenus à aborder dans cette partie textuelle, à savoir différents moyens linguistiques précédemment présentés : les marqueurs de topicalisation, de récapitulation, des marqueurs qui servent à exprimer une hypothèse.

Explication : Cette activité de production écrite permet à l'apprenant de mettre en pratique les moyens linguistiques acquis pour la structuration d'une conclusion.

Consignes :

L'enseignant peut faire travailler l'apprenant sur cet extrait à la suite des deux autres textes des Fiches pédagogiques 1 et 2. D'une part, l'apprenant aura compris quelques mots-clés de l'article au fil de sa lecture. D'autre part, sur le plan discursif, l'apprentissage des MD permet à l'apprenant de prendre conscience de la place de ces derniers dans la structuration de l'énoncé. En particulier, l'apprenant peut imaginer comment évoluent les idées dans le texte, comment sont construites les parties textuelles.

Nous regroupons les idées essentielles de cet extrait dans des blocs comme dans la Figure 6-10 : Schéma d'analyse - Fiche pédagogique 3. En ce qui concerne les MD, chaque groupe de MD correspond à un code de couleur : les marqueurs métadiscursifs sont soulignés en vert, les argumentatifs en jaune, les séquences à fonction modale en violet, et les routines sémantico-sémantique en bleu foncé. Dans cet extrait, les MD sont utilisés pour relever la différence entre les résultats de l'auteur avec d'autres travaux sur *l'acquisition du prétérit anglais par des apprenants français*.

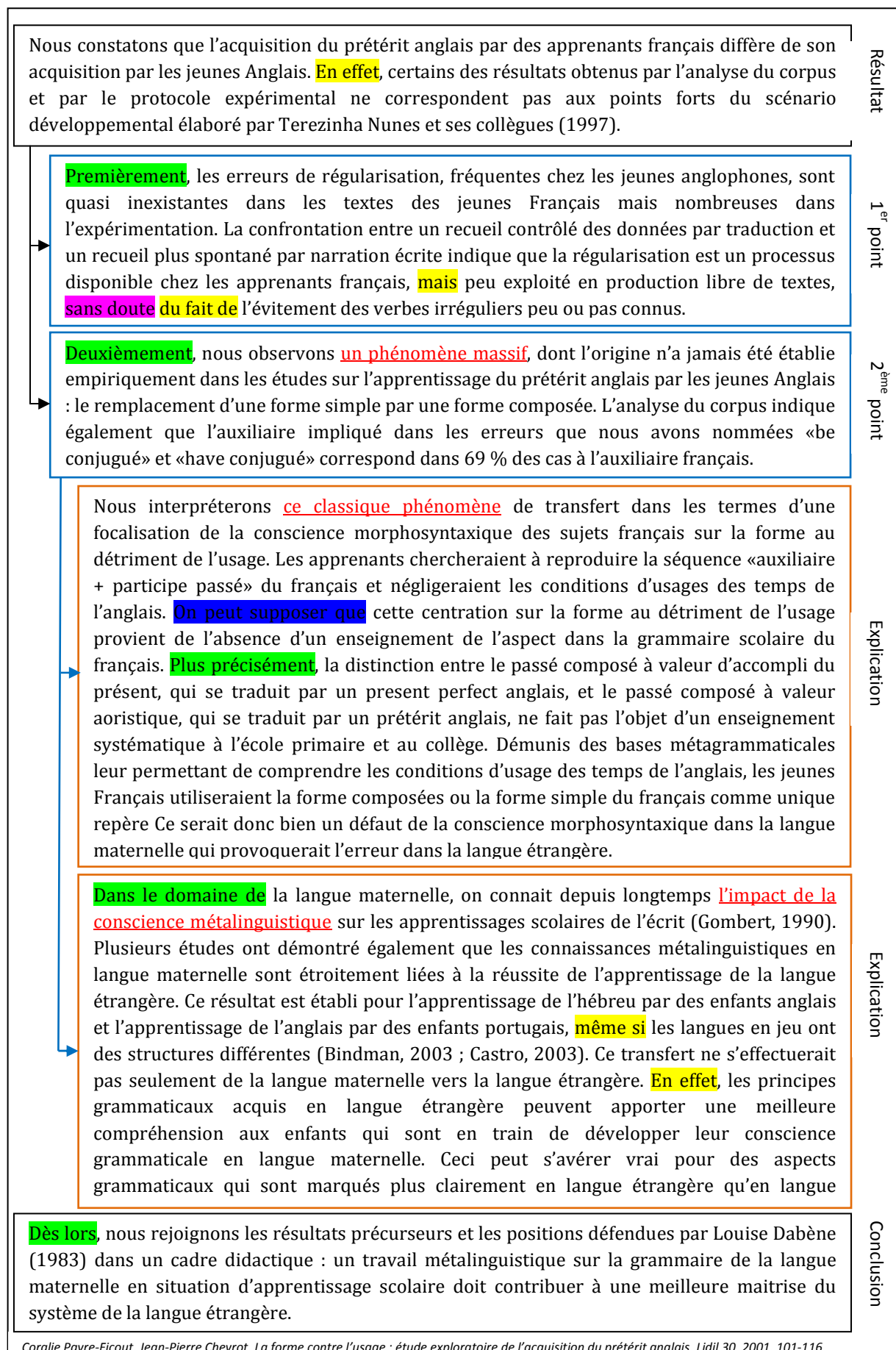


Figure 6-10 : Schéma d'analyse - Fiche pédagogique 3

Fiche pédagogique 4

Cet extrait correspond à la partie d'évaluation critique dans un article en traitement automatique des langues (TAL).

Ces trois expérimentations nous ont permis d'évaluer le modèle de représentation du texte, le langage SEXTANT et les interfaces de navigation qui sont proposées aux utilisateurs. En ce qui concerne le modèle de représentation du texte, les résultats confirment nos hypothèses : les différentes relations convoquées, syntagmatiques, syntaxiques, sémantiques et discursives ont pu être facilement représentées tout en conservant une bonne lisibilité. Il convient néanmoins de pondérer ce constat en remarquant que ces trois expérimentations ne convoquent pas de représentations intratextuelles complexes. En ce qui concerne le langage SEXTANT, les résultats sont là aussi satisfaisants, puisque toutes les opérations de navigation demandées par les experts ont pu être exprimées. Néanmoins, certaines limitations du langage de conditions sur lequel s'appuie le langage SEXTANT sont apparues. Notamment, il n'est pas possible d'exprimer une condition sur la position relative d'une UT cible dans le texte, autrement qu'en faisant référence à l'UT source. Si aucune des expérimentations n'a nécessité l'expression de telles conditions, il n'en reste pas moins que ceci pourrait constituer une limitation pour certains développements futurs. Quant à la gestion des interactions, le menu actuel de choix des opérations de navigation satisfait les besoins essentiels, mais sa position dans le texte ne peut pas être paramétrée, il se positionne toujours sur l'UT source, ce qui rend la lecture quelquefois difficile.

Ces expérimentations nécessitent l'expression d'opérations de navigation qui convoquent des expertises très différentes. L'expérimentation NaviLire a bénéficié de la longue expérience pédagogique et des réflexions théoriques de L. Lundquist sur l'enseignement de la linguistique textuelle. Nous disposons ainsi d'une documentation abondante (exemples d'exercices sur support imprimé, ouvrages publiés) que nous avons aisément transcrite sous la forme d'opérations de navigation. De même, les structures textuelles ont été balisées (manuellement) par L. Lundquist à partir de ces mêmes exercices. Pour l'expérimentation sur le résumé, nous avons déjà mentionné les travaux sur lesquels nous nous sommes appuyés et il faut simplement remarquer que les structures textuelles sont repérées automatiquement par un système dédié au filtrage d'informations. En revanche, les opérations de navigation, même si elles sont issues de divers travaux ont été spécifiées empiriquement et une évaluation en usage reste à faire. Enfin, l'expérimentation sur la navigation dans le roman « Madame Bovary » est la plus prospective mais elle a le mérite de montrer que des relations de types syntaxiques et ontologiques peuvent être représentées dans le formalisme proposé par NaviTexte.

Javier Couto, Jean-Luc Minel, Navigation textuelle : représentation des textes et des connaissances. Traitement Automatique des Langues, 47-2, 2006.

➤ **Niveau : B2/C1**

➤ **Objectifs :**

- Marqueurs de topicalisation
- Effectuer une évaluation des résultats

L'enseignant peut se référer à la Figure 6-11 : Schéma d'analyse - Fiche pédagogique 4 pour une représentation en blocs des idées.

➤ **1^{re} étape : Observation + repérage des MD**

Approche : inductive. Stratégie de lecture : *bottom-up*.

Activités : L'apprenant doit observer le texte sans les MD et tenter de préciser ce qui manque. Ensuite, l'enseignant demande de comparer le texte sans MD avec le texte original pour aider l'apprenant à repérer et prendre conscience du rôle de structuration des MD.

Explication : Cette activité se trouve en première phase d'apprentissage. Elle est une activité de découverte qui sert à repérer les MD. Ensuite, il conviendra de sensibiliser l'apprenant à l'importance des MD dans la construction d'un discours cohérent et structuré.

➤ **2^e étape : Classification des MD**

Approche : onomasiologique. Stratégie de lecture : *bottom-up*.

Activités : Après une étape de repérage des MD (davantage sémasiologique), l'enseignant demande à l'apprenant de regrouper les MD selon leur fonction rhétorique (donc classement onomasiologique) : marqueurs de topicalisation (*en ce qui concerne, quant à*), marqueurs de concession (*néanmoins, même si, en revanche*), marqueur de causalité (*puisque*), marqueurs de comparaison (*de même*).

Explication : Cet extrait est caractérisé par la fréquence des marqueurs de topicalisation. Ces marqueurs servent à introduire un nouveau contenu ou une nouvelle thématique et occupent une position charnière entre les idées. L'utilisation des MD permet en fait un repérage rapide des idées et facilite la lecture, ce qui constitue l'un des objectifs majeurs pour les apprenants.

➤ 3^e étape : Relever les idées essentielles

Approche : inductive onomasiologique. Stratégie de lecture *top-down*

Activités : L'enseignant pose des questions pour que l'apprenant prenne conscience de l'ordre des idées grâce à l'utilisation des marqueurs de topicalisation.

- *Quelle est la place de cet extrait dans un article scientifique ?*
- *Quelles sont les expérimentations effectuées dans ce travail ?*
- *Quel est l'objectif de ces expérimentations ?*
- *Sur quelle théorie s'effectuent ces expérimentations ?*

Après avoir relevé les idées essentielles du texte, l'enseignant peut demander à l'apprenant de les classer selon leur ordre d'apparition et de vérifier comment elles sont linguistiquement introduites.

Explication : Les MD contribuent à segmenter le texte en plusieurs constituants et en plusieurs idées et servent donc à faciliter la lecture. Cette activité sert à développer la stratégie de lecture *top-down*.

➤ 4^e étape : Collocations V + *hypothèse*

Approche : onomasiologique + inductive, aspect lexical.

Activités : L'enseignant fait observer à l'apprenant les concordances d'*hypothèse*.

Outre les marqueurs discursifs, l'apprenant dispose d'autres moyens exprimer l'hypothèse, essentiellement les collocations dont la structure syntaxique est de Verbe + *hypothèse*. Il existe dans le corpus Scientext des grammaires prédéfinies par des membres de notre équipe, l'apprenant peut accéder à de nombreux exemples pour exprimer l'hypothèse.

▼ N°	▼ Contexte gauche: 10 mots	▼ Occurrence:	▼ Contexte droit: 10 mots	▼ Réf. texte
✓ 1	L'	hypothèse émise	pour explorer les variations dans les usages adultes est la	[lin-com-14-body]
✓ 2	point de vue de sa stratification sociale on peut	émettre l' hypothèse	que les parents donnent la priorité à l' apprentissage de	[lin-com-14-body]
✓ 3	Von Hippel , chercheur de longue date dans le domaine de l' innovation économique ,	fait l' hypothèse	que nous sommes entrés dans une ère de " démocratisation	[sed-com-16-body]
✓ 4	de régulation de l' activité des gènes , et	formulent l' hypothèse	selon laquelle des mécanismes analogues expliquent les variations d' activité	[bio-art-18-introduction]
✓ 5	bactéries , des modifications possibles de l' ADN , l'	hypothèse selon laquelle la méthylation de l' ADN pourrait participer au contrôle de l' expression des gènes est proposée	simultanément par Arthur Riggs 5 et Robin Holliday 6 en 1975	[bio-art-18-introduction]
✓ 6	Ils	émettent l' hypothèse	qu' elle soit due à une accumulation de structures tourbillonnaires	[mec-the-19-body]
✓ 7	pas tous aux mêmes attributs de la dissertation et nous	faisons l' hypothèse	que cette ambiguïté engendre le risque d' enseigner une vision	[sed-com-28-body]
✓ 8	Ainsi , comme nous en	faisons l' hypothèse	, le sens de la notion de " problématique "	[sed-com-28-conclusion]

Explication : Cette activité s'intègre dans une phase d'appropriation des connaissances. Elle a pour objectif de donner plus de moyens lexicaux à l'apprenant quand ils rédigent. Ceci a été développé dans FULS (Cavalla, à paraître).

➤ **5^e étape : Production écrite**

Approche : rhétorique

Activités : L'apprenant rédige un texte d'une dizaine de lignes pour développer une idée sur son sujet de recherche en utilisant les marqueurs de topicalisation.

Explication : L'exercice de production écrite permet à l'apprenant de mobiliser les connaissances acquises pour atteindre l'objectif final pour la rédaction de son rapport de stage ou son mémoire.

Dans la Figure 6-11 : Schéma d'analyse - Fiche pédagogique 4, nous essayons de représenter, en les regroupant dans des cadres enchâssés, le rapport existant entre les idées développées par les auteurs. La lecture du texte peut se faire en plusieurs étapes. La première étape s'appuie sur les marqueurs de topicalisation pour avoir une vue globale de la structure du texte et sert à segmenter le texte en constituants. La deuxième étape sert à traiter le contenu sémantique dans chaque constituant. Dans cet extrait, les auteurs développent deux idées principales, à savoir : le rôle de trois expérimentations dans l'évaluation des outils, du langage, des interfaces de navigation et leurs fondements théoriques.

Ces trois expérimentations nous ont permis d'évaluer le modèle de représentation du texte, le langage SEXTANT et les interfaces de navigation qui sont proposées aux utilisateurs. **En ce qui concerne** le modèle de représentation du texte, les résultats confirment nos hypothèses : les différentes relations convoquées, syntagmatiques, syntaxiques, sémantiques et discursives ont pu être facilement représentées tout en conservant une bonne lisibilité. Il convient **néanmoins** de pondérer ce constat en remarquant que ces trois expérimentations ne convoquent pas de représentations intratextuelles complexes. **En ce qui concerne** le langage SEXTANT, les résultats sont là aussi satisfaisants, **puisque** toutes les opérations de navigation demandées par les experts ont pu être exprimées. **Néanmoins**, certaines limitations du langage de conditions sur lequel s'appuie le langage SEXTANT sont apparues. Notamment, il n'est pas possible d'exprimer une condition sur la position relative d'une UT cible dans le texte, autrement qu'en faisant référence à l'UT source. Si aucune des expérimentations n'a nécessité l'expression de telles conditions, il n'en reste pas moins que ceci pourrait constituer une limitation pour certains développements futurs.

Quant à la gestion des interactions, le menu actuel de choix des opérations de navigation satisfait les besoins essentiels, mais sa position dans le texte ne peut pas être paramétrée, il se positionne toujours sur l'UT source, ce qui rend la lecture quelquefois difficile.

Ces expérimentations nécessitent l'expression d'opérations de navigation qui convoquent des expertises très différentes. L'expérimentation NaviLire a bénéficié de la longue expérience pédagogique et des réflexions théoriques de L. Lundquist sur l'enseignement de la linguistique textuelle. Nous disposons ainsi d'une documentation abondante (exemples d'exercices sur support imprimé, ouvrages publiés) que nous avons aisément transcrite sous la forme d'opérations de navigation. **De même**, les structures textuelles ont été balisées (manuellement) par L. Lundquist à partir de ces mêmes exercices. Pour l'expérimentation sur le résumé, nous avons déjà mentionné les travaux sur lesquels nous nous sommes appuyés et il faut simplement remarquer que les structures textuelles sont repérées automatiquement par un système dédié au filtrage d'informations. **En revanche**, les opérations de navigation, même si elles sont issues de divers travaux ont été spécifiées empiriquement et une évaluation en usage reste à faire. **Enfin**, l'expérimentation sur la navigation dans le roman « Madame Bovary » est la plus prospective mais elle a le mérite de montrer que des relations de types syntaxiques et ontologiques peuvent être représentées dans le formalisme proposé par NaviTexte.

Javier Couto, Jean-Luc Minel, Navigation textuelle : représentation des textes et des connaissances. Traitement Automatique des Langues, 47-2, 2006.



Figure 6-11 : Schéma d'analyse - Fiche pédagogique 4

Conclusion

Nous allons conclure cette partie sur des réflexions didactiques sur l'enseignement/apprentissage des MD en général et des marqueurs de reformulation en particulier. Nous développerons trois points principaux, à savoir l'importance des descriptions linguistiques pour les activités didactiques, l'utilité d'un travail sur corpus et l'association des approches d'enseignement. Premièrement, nous confirmons l'importance des descriptions linguistiques menées dans le chapitre 5 pour la mise en place des activités scientifiques. Ces descriptions ont évidemment besoin d'être didactisées pour être utilisées par les apprenants. Elles permettent de sensibiliser l'enseignant et les apprenants aux spécificités d'usage de chaque MD. L'enseignant peut s'inspirer de ces données pour les adapter dans un programme d'apprentissage en fonction du niveau et de la progression des apprenants. Deuxièmement, le travail sur corpus peut être très bénéfique pour l'apprentissage des unités lexicales. Pour un apprentissage efficace, les enseignants de FLE devraient être les premiers à constater la plus-value de l'apprentissage sur corpus et des concordanciers, et prêts à l'intégrer dans leurs cours de langue.

Enfin, l'association des quatre approches présentées s'avère utile. En fait, les approches onomasiologique et sémasiologique sont introduites dans les méthodes de FLE et elles sont mêlées. En classe de langue, l'enseignant alterne les approches en suivant la progression des apprenants. Nous avons vu que l'approche inductive permet de développer chez les apprenants la compétence d'observation. Cependant, à une étape de consolidation, une approche déductive semble aussi nécessaire. L'approche onomasiologique sert à englober les fonctions principales des MD, ce qui est indispensable pour les apprenants pour développer la compétence rédactionnelle. L'approche sémasiologique, quant à elle, est nécessaire pour l'enseignant qui souhaite revenir sur une unité linguistique (ou un MD) pour mieux l'approfondir. Elle est également utile à l'apprenant afin de fixer dans sa mémoire ses observations pour une exploitation ultérieure.

Conclusion générale

Nous allons retracer, dans les lignes qui suivent, les principaux résultats de ce travail, ainsi que les perspectives de recherche que nous envisageons à l'issue de cette thèse.

Dans ce travail, nous nous étions fixé deux objectifs d'ordre linguistique et didactique. En ce qui concerne le premier objectif, d'ordre linguistique, l'étude du corpus d'articles scientifiques nous a permis de mettre en place une typologie des séquences lexicalisées à fonction discursive en nous appuyant sur leurs fonctions dans une approche onomasiologique, ainsi que de relever leurs spécificités dans les écrits scientifiques en SHS. Certains marqueurs paraissent particulièrement représentatifs des disciplines de SHS comme les marqueurs de reformulation, d'exemplification ou de comparaison.

Le modèle d'analyse que nous avons mis en place pour modéliser les marqueurs de reformulation s'avère efficace pour traiter le phénomène des MD en associant leurs paramètres syntaxiques et sémantiques (portée, position, valeurs sémantiques). À notre connaissance, il s'agit de la première modélisation des marqueurs de reformulation dans les écrits scientifiques. Un tel modèle peut être adaptable à l'analyse d'autres unités lexicales.

Pour notre deuxième objectif, d'ordre didactique, nous avons montré comment l'analyse linguistique basée sur corpus peut contribuer à la conception d'activités didactiques. La transposition didactique des connaissances tirées des descriptions linguistiques, qui n'est pas une étape aisée, est indispensable pour une adaptation adéquate à notre public universitaire. En effet, la modélisation des MD nous a permis de relever leurs régularités syntaxiques et sémantiques, ce qui nous permet d'aborder tous les aspects linguistiques de ces éléments phraséologiques dans l'enseignement.

Au plan didactique, nous proposons une entrée méthodologique qui prend en compte différentes approches pour enseigner des phénomènes linguistiques à l'aide du corpus. Les enseignants y trouvent des considérations méthodologiques pour aborder les MD grâce au corpus. L'enseignement de ces éléments est envisagé du point de vue métalinguistique, en

soulignant l'acquisition des connaissances d'une manière consciente pour les apprenants. Une maîtrise métalinguistique se construit en plusieurs étapes et doit commencer par la prise de conscience des régularités linguistiques. Cette étape métalinguistique nous semble très importante pour améliorer des compétences rédactionnelles des apprenants.

En outre, nous avons montré que l'enseignement/apprentissage des marqueurs de reformulation ne s'arrête pas à la dimension lexicale. Pour un enseignement systématique de ces unités lexicales, il est nécessaire de les intégrer dans une chaîne d'écriture et initier les apprenants aux structures discursives et aux spécificités des écrits scientifiques (« routines sémantico-rhétoriques », etc.).

La linguistique de corpus apporte des pistes d'explorations intéressantes, particulièrement pour la modélisation des MD et pour des applications didactiques. Le fait de travailler sur un corpus annoté offre de nombreux avantages. Le corpus sert à alimenter nos activités didactiques. Grâce aux techniques du TAL, nous avons mis en place une base de données de séquences lexicalisées à fonction discursive qui fonctionne comme un dictionnaire sur corpus fondé sur la typologie antérieurement conçue. En particulier, le corpus constitue un outil pertinent aussi bien pour les enseignants natifs ou non natifs que pour les étudiants. Les enseignants non natifs ne peuvent pas recourir à leur intuition et peuvent aisément trouver des explications pour des phénomènes linguistiques à partir des observations sur corpus.

Ce travail est, à notre connaissance, une première expérimentation d'intégration d'un corpus d'écrits scientifiques en classe de langue pour enseigner les MD. Il n'a pas pour but de concevoir un outil d'aide à l'écriture scientifique « fini » ni de monter des séquences didactiques. Nous souhaitons développer des réflexions didactiques au service de l'enseignement/apprentissage d'un lexique spécifique à partir d'observation sur corpus. Ces réflexions permettront d'élaborer une méthodologie qui sera exploitée par la suite à l'enseignement/apprentissage d'autres unités lexicales, dans d'autres domaines.

Ce travail propose une autre manière d'aborder la dimension linguistique dans les programmes du FOU. L'enseignement n'a rarement qu'un seul objectif. Les activités que nous avons décrites montrent que l'enseignement des MD peut être traité sous plusieurs angles, aussi bien lexical que structurel. Certes, nous ne pouvons pas tout faire avec les MD, mais l'enseignement/apprentissage de ces unités lexicales représente une première étape qui permet de développer des compétences rédactionnelles et discursives chez les apprenants. L'apprentissage des MD ou d'autres éléments linguistiques doit commencer par une étape de prise de conscience, en passant par l'analyse et l'entraînement, pour arriver à la réutilisation. Un tel enseignement a pour objectif de sensibiliser les apprenants aux moyens servant à

organiser le texte et à construire un discours reconnu par la communauté des chercheurs, ainsi que de les aider à mieux comprendre les règles conventionnelles des écrits scientifiques.

Les considérations méthodologiques que nous avons évoquées dans ce travail pourraient être intégrées dans des manuels destinés à un public universitaire. Elles peuvent s'adresser aussi bien aux étudiants locuteurs non natifs du français qu'aux étudiants locuteurs natifs. Les étudiants locuteurs non natifs du français ont besoin d'avoir à leur disposition plus de moyens pour la pratique de l'écriture scientifique. Ces considérations peuvent également être utiles à notre avis aux locuteurs natifs, qui doivent eux aussi s'habituer et s'approprier des règles d'un nouveau genre d'écrit qu'ils n'ont pas l'occasion de pratiquer antérieurement dans leur cursus scolaire.

Perspectives et limites de recherche

Ce travail présente bien évidemment encore quelques limites. La première concerne le plan didactique : il nous manque d'avoir pu appuyer nos analyses sur une expérimentation que nous comptons mener prochainement dans une université au Vietnam, dans des cours de FOU, introduits au Vietnam depuis quelques années dans les programmes des Filières Universitaires Francophones (FUF). Cette expérimentation s'avère indispensable, et nous permettra d'analyser les pratiques d'enseignement/apprentissage dans le contexte du Vietnam, ainsi que de tester la mise en place des activités. Elle nous aidera également à réajuster et à mieux adapter nos propositions didactiques pour le contexte d'enseignement/apprentissage du pays. Peut-être pourrons-nous mieux définir la place de l'enseignant face à l'introduction d'un corpus en classe de langue et de juger de l'utilité, dans l'enseignement, des phénomènes linguistiques à l'aide de corpus.

La deuxième limite est liée à la taille du corpus. Certains marqueurs sont exclus en raison de leur faible fréquence dans notre corpus. En travaillant sur un corpus plus volumineux, nous pourrions enrichir notre typologie des séquences lexicalisées à fonction discursive. Enfin, nous rencontrons une autre difficulté liée aux droits d'auteur, puisque nous ne pouvons accéder qu'à une partie du contexte pour certains textes. En travaillant sur les aspects discursifs, il est indispensable d'avoir un contexte relativement large pour mieux comprendre comment les idées se développent et comment un MD est utilisé.

Ce travail de thèse n'est pas un aboutissement final et permet d'envisager plusieurs perspectives pour des travaux ultérieurs.

Au plan linguistique, outre les unités polylexicales de cette étude, nous souhaitons également analyser le comportement syntaxique et sémantique des MD monolexicaux. Il est intéressant de comparer le fonctionnement des marqueurs monolexicaux et polylexicaux de même catégorie afin de relever leurs régularités syntaxiques et sémantiques, et de mieux comprendre le comportement des polylexicaux. Nous remarquons que certaines opérations s'effectuent essentiellement avec des marqueurs polylexicaux comme la reformulation, l'exemplification et la comparaison. L'analyse de la chaîne syntagmatique associée aux MD peut être une des pistes à envisager.

Une autre piste de recherche à envisager se situe dans le cadre de l'analyse contrastive des MD en français avec d'autres langues, en vietnamien par exemple (*Cf.* Chapitre 6 Corpus de mémoires d'apprenants vietnamophones). Cette mise en comparaison permettrait de relever les ressemblances et aussi les écarts entre les deux langues. Elle servirait peut-être à expliquer les difficultés des étudiants vietnamiens dans la construction d'un discours scientifique argumenté.

Sur le plan didactique, nous avons pour objectif à plus long terme de concevoir un prototype d'outil d'aide à l'écriture scientifique. En nous inscrivant dans cette perspective, un apprentissage qui porte sur le genre de texte nous semble indispensable. Nous envisageons donc de tirer profit davantage des recherches approfondies sur le genre (Rinck, 2006, Poudat, 2006). Pour bien rédiger l'écriture scientifique, les apprenants doivent avant tout acquérir des règles textuelles, ainsi que des spécificités de chaque genre, les pratiques d'écritures et comprendre les normes requises au sein d'une communauté scientifique. Leurs connaissances sur le genre et les pratiques d'écriture de ce genre sont importantes et il serait important de les aborder avec les étudiants.

En outre, la modélisation des marqueurs de reformulation que nous avons effectuée dans cette thèse constitue une première étape pour repérer des propriétés sémantiques nécessaires pour d'autres applications lexicographiques. Au fil des analyses linguistiques, nous avons constaté que certains MD dans le TLFi ne sont parfois pas très clairement définis. Ils sont souvent définis dans un rapport synonymique avec d'autres MD, ce qui ne permet pas de faire ressortir la spécificité d'emploi de chacun.

Par ailleurs, nous envisageons une autre piste de travail concernant la constitution d'un corpus plus important d'apprenants locuteurs non natifs du français. Celui-ci doit être comparé avec le corpus d'experts à l'instar de Paquot (2010). Cette comparaison de deux genres représente pour nous une perspective didactique intéressante.

Enfin, nous souhaitons procéder à l'annotation sémantique des MD sur les corpus, avec la perspective d'élargir les champs d'analyse aux monolexicaux. Le travail d'annotation des MD nous semble très intéressant, car il permet aux apprenants de repérer rapidement les structures discursives du texte et de leur donner une vision d'ensemble. Par exemple, nous constatons que les marqueurs d'organisation textuelle comme les marqueurs de topicalisation ou marqueurs d'énumération permettent de repérer rapidement la thématique introduite par ces marqueurs.

Pour finir, nous souhaitons rappeler que l'enseignement/apprentissage des MD constitue un des éléments essentiels à aborder pour la pratique d'écriture scientifique, et représente pour nous un premier pas vers la construction d'un discours cohérent permettant de véhiculer la pensée scientifique.

Bibliographie

- Adam, J.-M. (1989). Aspects de la structuration du texte descriptif: les marqueurs d'énumération et de reformulation. *Langue française*, (81), 59-98.
- Allouche, V., & Maurer, B. (2011). *L'écrit en FLE : travail du style et maîtrise de la langue*. Montpellier : Presses universitaires de la Méditerranée.
- Altenberg, B. (1998). On the phraseology of spoken English: the evidence of recurrent word-combinations. In Cowie, Anthony P. (Éd.), *Phraseology: Theory, Analysis, and Applications*. Oxford : Oxford University Press, 101-122.
- Anctil, Dominic. (2005). *Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes lexicaux et analyse des stratégies de résolution de problèmes lexicaux*. Maîtrise en didactique. Montréal : Université de Montréal.
- Anctil, D. (2006). Caractérisation de la notion d'erreur lexicale et description des problèmes lexicaux d'étudiants universitaires. In *Autour des langues et du langage : perspectives multidisciplinaires, Colloque international des étudiants-chercheurs en linguistique et didactique des langues*, Grenoble, 4-7 juillet 2006. Grenoble : Presses universitaires de Grenoble, 349-356.
- Anscombre, J.-C. (1983). Pour autant, pourtant (et comment) : à petites causes grands effets. *Cahiers de linguistique française*, (5), 37-84.
- Aston, G. (2002). The Learner as Corpus Designer. *Language and Computers*, 42(1), 9-25.
- Auchlin, A. (1981). Marqueurs de structuration de la conversation et complétude. *Cahiers de linguistique française*, (1), 141-160.
- Bard, A. (2007). *Constitution d'un corpus d'écrits universitaires d'étudiants étrangers* (Mémoire de Master 1 FLE Professionnel). Université Stendhal Grenoble 3, Grenoble.
- Berthelot, J.-M. (Éd.). (2003). *Figures du texte scientifique*. Paris : Presses universitaires de France.
- Biber, D. (2006). *University language : a corpus-based study of spoken and written registers*. Amsterdam, Philadelphia : John Benjamins.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *English for Specific Purposes*, (23), 397-423.
- Biber, D. & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In H. Hasselgard & S. Oksefjell (Éd.) : *Out of Corpora: Studies in honor of Stig Johansson*. Amsterdam : Rodopi, 181-189.
- Biber, D., Conrad, S. & Cortes, V. (2003). Lexical bundles in speech and writing : An initial taxonomy. In A. Wilson, P. Rayson & T. McEnery (Éd.) : *Corpus Linguistics by the Lune*. Frankfurt/Main : Peter Lang, 71-93.
- Blondel, C. (2012). *Quelle formation en grammaire et en orthographe pour des étudiants de licence : de l'étude des besoins aux propositions didactiques* (Mémoire Master 2). Université Stendhal - Grenoble 3, Grenoble. Consulté le 22 octobre 2014 : http://dumas.ccsd.cnrs.fr/docs/00/75/14/87/PDF/MA_moire-M2R-Blondel.pdf

- Blumenthal, P. (2007). Sciences de l'Homme vs sciences exactes : combinatoire des mots dans la vulgarisation scientifique. *Revue Française de Linguistique Appliquée*, XII(2), 15-28.
- Boch, F. (2013). Former les doctorants à l'écriture de la thèse en exploitant les études descriptives de l'écrit scientifique. *Linguagem em (Dis)curso*, 13(3), 543-568.
- Bonnot, C. (2001). La portée des mots du discours : essai de définition (sur l'exemple du russe moderne). *Cahiers de linguistique de l'INALCO*, (4), 9-30.
- Boulton, A. (2008). Esprit de corpus : promouvoir l'exploitation de corpus en apprentissage des langues. *Texte et Corpus*, (3), 37-46.
- Boulton, A. (2010). Data-Driven Learning: On Paper, in Practice. In T. Harris, & M. Moreno Jaén, (Éd.), *Corpus linguistics in language teaching*. Bern ; Berlin ; Bruxelles [etc.] : Peter Lang, 17-52.
- Boulton, A. (2011). What data for data-driven learning? In Actes *European Association for Computer-Assisted Language Learning (EUROCALL)*, Nottingham, 23-27.
- Boulton, A., & Tyne, H. (2014). *Des documents authentiques aux corpus. Démarches pour l'apprentissage des langues*. Paris : Didier.
- Bourigault, D. (2007). *Un analyseur syntaxique opérationnel : SYNTAX* (mémoire de HDR). Toulouse.
- Burger, H. (1998). *Phraseologie: eine Einführung am Beispiel des Deutschen*. Berlin : Erich Schmidt.
- Carras, C., & Cavalla, C. (2007). Développement d'outils pour le suivi linguistique des apprenants dans les universités françaises. Présenté à *L'ouverture des filières universitaires francophones aux étudiants étrangers : enjeux politiques, implications didactiques, culturelles et institutionnelles*, ADCUEF, CLA-Université de Franche-Comté.
- Carras, C., Gewirtz, O., & Tolas, J. (2014). *Réussir ses études d'ingénieur en français*. Grenoble : Presses Universitaires de Grenoble.
- Cavalla, C. (à paraître). Collocations transdisciplinaires dans les écrits de doctorants FLS/FLE. In F. Sitri & F. Rinck (Éd.), *Former à l'écrit universitaire*. Paris : Linx.
- Cavalla, C. (2014). Collocations transdisciplinaires : réflexion pour l'enseignement. In M.-I. González Rey (Éd.), *Outils et méthode d'apprentissage en phraséodidactique*. Fernelmont, Belgique : E.M.E, 151-169.
- Cavalla, C. (2007). Réflexion pour l'aide à l'écrit universitaire auprès des étudiants étrangers entrant en Master et Doctorat. In J. Goes & J.-M. Mangiante (Éd.). Présenté à *L'accueil des étudiants étrangers dans les universités francophones*, Arras : Artois Presses Université, 37-48.
- Cavalla, C. (2008). Propositions didactiques pour l'enseignement d'éléments phraséologiques en FLE. In D. Van Raemdonck & M.-E. Damar (Éd.), *Enseigner les structures langagières en FLE*. Consulté le 22 octobre 2014 : <http://tmaes.no-ip.org/ME/>
- Cavalla, C. (2009). La phraséologie en classe de FLE. *Les Langues Modernes*, (1). Consulté le 24 octobre 2014 : <http://www.aplv-languesmodernes.org/spip.php?article2292>

- Cavalla, C., & Loiseau, M. (2013). Scientext comme corpus pour l'enseignement. In A. Tutin & F. Grossmann (Éd.), *L'écrit scientifique : du lexique au discours. Autour de scientext*. Rennes : Presses Universitaires de Rennes, 163-180.
- Chambers, A. (2010). L'apprentissage de l'écriture en langue seconde à l'aide d'un corpus spécialisé. *Revue française de linguistique appliquée*, XV, 9-20.
- Charolles, M. (1997). L'encadrement du discours : univers, champs, domaines et espaces. *Cahier de Recherche Linguistique, LANDISCO*, (6), 1-73.
- Charolles, M. (2003). De la topicalité des adverbiaux détachés en tête de phrase. *Travaux de linguistique*, (47), 11-50.
- Charolles, M., & Coltier, D. (1986). Le contrôle de la compréhension dans une activité rédactionnelle : Eléments pour l'analyse des reformulations paraphrastiques. *Pratiques*, (49), 51-68.
- Charolles, M., & Péry-Woodley, M.-P. (2005). *Les adverbiaux cadratifs*. Paris : Larousse.
- Charolles, M., & Vigier, D. (2005). Les adverbiaux en position préverbale : portée cadrative et organisation des discours. *Langue française*, (148), 9-30.
- Charolles, M., & Prévost, S. (Éd.) (2003). Adverbiaux et topiques. *Travaux de linguistique*, (47).
- Chetouani, L. (1988). *Vocabulaire général d'enseignement scientifique (VGES) : recherche sur le vocabulaire de la communication scientifique orale en français dans les classes de lycée en Algérie* (Thèse de doctorat). Université de Paris XIII, Villetaneuse.
- Chetouani, L. (1997). *Vocabulaire général d'enseignement scientifique, VGES*. Editions L'Harmattan.
- Cislaru, G., Claudel, C., & Vlad, M. (2009). *L'écrit universitaire en pratique : applications et exercices autocorrectifs, rédaction d'un mémoire, présentation PowerPoint*. Bruxelles : De Boeck.
- Colletta, J.-M. (2004). *Le développement de la parole chez l'enfant âgé de 6 et 11 ans : corps, langage et cognition*. Sprimont : Editions Mardaga.
- Combettes, B. (1998). *Les constructions détachées en français*. Gap : Ophrys.
- Conseil de l'Europe. (2001). *Cadre européen commun de référence pour les langues*. Paris : Didier.
- Coxhead, A. (2000). « A New Academic Word List ». *TESOL Quarterly*, 34(2), 213-238.
- Coxhead, A. (2002). The Academic Word List: A Corpus-based Word List for Academic Purposes. In *Teaching and Language Corpora (TALC) 2000 Conference Proceedings*. Atlanta : Rodopi.
- Crompton, P. (2006). The effect of position on the discourse scope of adverbial. *Text & Talk*, 26(3), 245-279.
- Cuenca, M. J., & Bach, C. (2007). Contrasting the form and use of reformulation markers. *Discourse Studies*, 9(2), 149-175.
- Danell, K. J. (1998). La portée comme phénomène linguistique. *Le français moderne*, LXVI(1), 1-26.
- Danjou-Flaux, N. (1980). A propos du de fait, en fait, en effet, effectivement. *Le français moderne*, 48(1-2), 110-139.

- Da Sylva, L. (2010). Extraction semi-automatique servant de base pour l'indexation automatique. In *Actes TALN 2010*, Montréal.
- De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures (BELL), New Series*, (2), 225-246.
- Delahaie, J. (2011). Les marqueurs discursifs, un objet d'enseignement pertinent pour les étudiants Erasmus? *Etudes de linguistique appliquée*, (162), 153-164.
- Delahaie, J., & Flament-Boistrancourt, D. (2013). Corpus et enseignants de français L2 en Flandre : une histoire réussie. *Bulletin suisse de linguistique appliquée (VALS-ASLA)*, (97), 77-96.
- Diep, Kien Vu. *Analyse linguistique de la cohérence dans l'explication scientifique (le cas du discours agronomique)* (Thèse de doctorat). Université de Rouen, 2008.
- Dostie, G., & Pusch, C. D. (2007). Présentation. *Langue Française*, (154), 3-12.
- Drouin, P. (2007). Identification automatique du lexique scientifique transdisciplinaire. *Revue française de linguistique appliquée*, XII(2), 45-64.
- Eckenschwiller, M. (1994). *L'écrit universitaire*. Paris : les éd. d'Organisation.
- Eurin Balmet, S., & Henao de Legge, M. (1992). *Pratiques du français scientifique*. Vanves : Hachette.
- Falaise, A. (2014). Exploitation linguistique de corpus arborés d'écrits scientifiques à l'aide du logiciel ScienQuest. In A. Tutin & O. Kraif (Éd.), *L'écrit scientifique : du lexique au discours. Autour de scientext*. Rennes : Presse Universitaires de Rennes, 123-141.
- Falaise, A., Tutin, A., & Kraif, O. (2011). Une interface pour l'exploitation de corpus arborés par des non informaticiens : la plate-forme ScienQuest du projet Scientext. *TAL*, 52(3), 241-246.
- Falaise, A., Tutin, A., Kraif, O., & Rouquet, D. (2012). ScienQuest: a treebank exploitation tool for non NLP-specialists. In *Proceedings of COLING 2012*. Mumbai, India, 131-140.
- Fernandez, M. M. J. (1994). *Les particules énonciatives*. Paris : Presses universitaires de France.
- Fligelstolle, S. (1993). Some reflections on the question of teaching, from a corpus linguistics perspective. *ICAME journal*, 17, 97-109.
- Flower, L., & Hayes, J. R. (1980). The dynamics of composing : Making plans and juggling constraints. *Cognitive processes in writing*, (31), 31-50.
- Garnier, S., & Savage, A. D. (2011). *Rédiger un texte académique en français : règles grammaticales, règles d'usage, exercices d'entraînement corrigés*. Paris : Ophrys.
- Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching. In A. O'Keeffe & M. McCarthy (Éd.), *The Routledge handbook of corpus linguistics*. Routledge : 359-370.
- Gledhill, C. (2000). *Collocations in science writing*. Tübingen : Gunter Narr Verlag Tübingen.
- Glorieux, C., Pollet, M.-C., Malengreau, E., Wynsberghe, D., & Delforge, M. (2006). Elaborer des outils d'aide à la rédaction d'écrits scientifiques : dépasser les lieux communs pour s'ancrer dans des besoins spécifiques. In *Actes de la journée d'étude : Mons, FUCaM*. Bruxelles : Conseil interuniversitaire de la Communauté française, 98-122.

- Gougenheim, G., Michéa, R., Rivenc, P., & Sauvageot, A. (1964). *L'élaboration du français fondamental (premier degré): étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Paris : Didier.
- Granger, S., & Meunier, F. (2008). Phraseology in Language Learning and Teaching. Where to from here? In *Phraseology in Foreign Language Learning and Teaching*. Benjamins : Amsterdam & Philadelphia, 247-252.
- Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Éd.), *Phraseology : An interdisciplinary perspective*. Amsterdam, Philadelphia : John Benjamins Publishing company, 27-49.
- Granger, S., & Paquot, M. (2009). *eLexicography in the 21st century: New challenges, new applications* (CECL.). Louvain-la-Neuve. Consulté le 22 octobre 2014 : http://www.uclouvain.be/cps/ucl/doc/cecl/documents/eLEX_book_abstracts_FINAL_New.pdf
- Granger, S., & Paquot, M. (2010). Customising a general EAP dictionary to meet learner needs. In S. Granger & M. Paquot (Éd.), *ELexicography in the 21st century: New Challenges, new applications*. Louvain-la-Neuve : Presses universitaires de Louvain, 87-96.
- Guimier, C. (1996). *Les adverbes du français : le cas des adverbes en -ment*. Gap ; Paris : Ophrys.
- Güllich, E., & Kotschi, T. (1983). Les marqueurs de reformulation paraphrastique. *Cahiers de linguistique français*, (5), 305-346.
- Habert, B., Nazarenko, A., & Salem, A. (1997). *Les linguistiques de corpus*. Paris : Masson & Armand Colin.
- Ho-Dac, L.-M., Fabre, C., Péry-Woodley, M.-P., Rebeyrolle, J., & Tanguy, L. (2012). An Empirical Approach to the Signalling of Enumerative Structures. *Discours. Revue de Linguistique, Psycholinguistique et Informatique*, (10). Consulté le 22 octobre 2014 : <http://discours.revues.org/8611>
- Howarth, P. (1998). The phraseology of learners' academic writing. In A. P. Cowie (Éd.), *Phraseology: theory, analysis and applications*. Oxford : Oxford University Press, 161-186.
- Huumo, T. (1996). A scoping hierarchy of locatives. *Cognitive linguistics*, (7-3), 265-299.
- Jacobi, D. (1999). *La communication scientifique*. Grenoble : Presses universitaires de Grenoble.
- Johns, T. (1991). Should you be persuaded : Two examples of data-driven learning. *English Language Research Journal*, (4), 1-16.
- Kamber, A. (2011). Contexte et sens: utilisation d'un corpus écrit dans l'enseignement / apprentissage du FLE. *Revue Tranel (Travaux neuchâtelois de linguistique)*, (55), 199-218.
- Kerbrat-Orecchioni, C. (2005). *Le discours en interaction*. Paris : Armand Colin.
- Khatchatourian, E. (2006). *Les mots du discours formés à partir des verbes dire/skazat' (en français et en russe)* (Thèse de doctorat). Université Paris 7, Paris.
- Khatchatourian, E. (2008). Les marqueurs de reformulation formés à partir du verbe dire. In M.-C. Le Bot, M. Shuwer, & E. Richard (Éd.), *La reformulation : marqueurs linguistiques, stratégies énonciatives*. Rennes : Presses universitaires de Rennes, 19-33.

- Kocourek, R. (1991). *La langue française de la technique et de la science : vers une linguistique de la langue savante* (2^e éd. augmentée, refondue et mise à jour avec une nouvelle bibl.). Wiesbaden : O. Brandstetter.
- Kraif, O. (2008). Comment allier la puissance du TAL et la simplicité d'utilisation ? l'exemple du concordancier bilingue ConcQuest. In *Actes des 9^{ème} Journées d'analyse statistique des données textuelles, JADT 2008*. Lyon : Presses universitaires de Lyon, 625-634.
- Laignelet, M. (2004). *Les titres et les cadres discours temporels* (DEA Sciences du langage). Université de Toulouse 2 - Le Mirail, Toulouse.
- Landure, C. (2013). Corpus et dictionnaires : prise de conscience linguistique chez des apprenants LANSAD. *Bulletin vals-asla*, (97), 59-76.
- Lasnier, F. (2000). *Réussir la formation par compétences*. Montréal : Guérin.
- Leech, G. (2005). Developing Linguistic Corpora: a Guide to Good Practice. In M. Wynne (Éd.), *Developing linguistic corpora: a guide to good practice* (Vol. 92). Oxford : Oxbrow Books, 17-29.
- Legallois, D., & Tutin, A. (2013). Présentation : Vers une extension du domaine de la phraséologie. *Langages*, 1(189), 3-25.
- Le, N. B. (2013). *L'utilisation des connecteurs argumentatifs dans les écrits universitaires des étudiants vietnamiens : le cas du discours géographie. Pour une perspective didactique* (Thèse de doctorat en sciences du langage Spécialité Français Langue Etrangère). Université de Grenoble, Grenoble.
- Lenepveu, V. (2008). En fin de compte, au bout du compte, tout compte fait : quand reformuler, c'est reconsidérer. In M.-C. Le Bot, M. Shuwer, & E. Richard (Éd.), *La reformulation : marqueurs linguistiques, stratégies énonciatives*. Rennes : Presses universitaires de Rennes, 113-126.
- Loffler-Laurian, A.-M. (1983). Typologie des discours scientifiques : deux approches. *Etudes de linguistique appliquée*, (51), 8-20.
- Lundquist, L. (2013). *Lire un texte académique en français*. Paris : Ophrys.
- Mangenot, F. (1997). Le multimédia dans l'enseignement des langues. In J. Crinon & C. Gautellier (Éd.), *Apprendre avec le multimédia, Où en est-on?*. Paris : Retz, 119-134.
- Mangiante, J.-M., & Parpette, C. (Éd.). (2010). *Faire des études supérieures en langue française*. Paris : Clé International.
- Mangiante, J.-M. & Parpette, C. (2011). *Le français sur objectif universitaire*. Grenoble : Presses universitaires de Grenoble.
- Mangiante, J.-M., & Parpette, C. (2012). Le Français sur Objectif Universitaire : de la maîtrise linguistique aux compétences universitaires. *Synergies Algérie*, (15), 147-166.
- Martinez, R., & Schmitt, N. (2012). A Phrasal Expressions List. *Applied Linguistics*, 33(3), 299-320.
- Montagnon, G. (2007). *Développement d'une séquence didactique pour l'utilisation des collocations dans les écrits universitaires* (Mémoire de Master 1 FLE Professionnel). Université Stendhal - Grenoble 3, Grenoble.

- Mroué, M. (2014). *Éléments pour une sensibilisation au positionnement scientifique à travers la phraséologie transdisciplinaire* (Thèse de doctorat). Université Grenoble Alpes, Grenoble.
- Muller, C. (1991). *La négation en français*. Genève : Droz.
- N'Da, P. (2007). *Méthodologie et guide pratique du mémoire de recherche et de la thèse de doctorat en lettres, arts, sciences humaines et sociales : informations, normes et recommandations universitaires, techniques et pratiques actuelles*. Paris : L'Harmattan.
- Nesselhauf, N. (Éd.). (2005). *Collocation in a Learner Corpus* (Vol. 14). Amsterdam / Philadelphia : John Benjamins Publishing Company.
- Osborne, J. (1994). La cohésion dans les productions écrites d'étudiants en anglais de spécialité : un problème culturel ? *ASp, la revue du GERAS*, (5-6), 205-216.
- Osborne, J. (2008). Phraseology effects as a trigger for errors in L2 English: The case of more advanced learners. In F. Meunier & S. Granger (Éd.), *Phraseology in Foreign Language Learning and Teaching*. Amsterdam / Philadelphia : John Benjamins Publishing Company, 67-84.
- Paillard, D., & Vu, T. N. (2012). *Inventaire raisonné des marqueurs discursifs du français. Description. Comparaison. Didactique*. Paris : AUF.
- Paquot, M. (2010). *Academic vocabulary in learner writing : from extraction to analysis*. London ; New-York : Continuum international publishing group.
- Paquot, M., & Bestgen, Y. (2009). Distinctive words in academic writing: A comparison of three statistical tests for keyword extraction. In A. H. Jucker, D. Schreier, & M. Hundt (Éd.), *Corpora: Pragmatics and Discourse*. Amsterdam : Rodopi, 247-269.
- Pecman, M. (2004). *Phraséologie contrastive anglais-français: analyse et traitement en vue de l'aide à la rédaction scientifique* (Thèse en Sciences du Langage). Université Sophia Antipolis, UFR Lettres, Arts et Sciences Humaines, Nice.
- Pecman, M. (2007). Approche onomasiologique de la langue scientifique générale. *Revue Française de Linguistique Appliquée*, XII(2), 79-96.
- Péry-Woodley, M.-P. (2005). Discours, corpus, traitements automatiques. *Sémantique et Corpus*, 177-210.
- Pétroff, A.-J. (1984). Sémiologie de la reformulation dans le discours scientifique et technique. *Langue Française*, (64), 53-67.
- Phal, A. (1971). *Vocabulaire général d'orientation scientifique*. Paris : Crédif.
- Picoche, J. (1992). *Précis de lexicologie française : l'étude et l'enseignement du vocabulaire* (Nouvelle éd. revue et mise à jour.). Paris : Nathan.
- Picoche, J. (1993). *Didactique du vocabulaire français*. Paris : Nathan.
- Polguère, A. (2008). *Lexicologie et sémantique lexicale : notions fondamentales* (Nouvelle édition revue et augmentée.). Montréal : Presses de l'Université de Montréal.
- Porhiel, S. (2005). Les séquences thématiques. *Langue française*, 148, 111-126.
- Poudat, C. (2006). *Etudes contrastive de l'article scientifique de revue linguistique dans une perspective d'analyse des genres* (Thèse en Sciences du Langage). Université d'Orléans, Orléans.

- Rastier, F. (2002). Enjeux épistémologique de la linguistique de corpus. In W. Geoffrey (Éd.), *Deuxièmes Journées de Linguistique de Corpus*, Lorient. Consulté à l'adresse http://www.revue-texto.net/Inedits/Rastier/Rastier_Enjeux.html
- Riabtseva, N. (1999). *English for scientific purposes: Guide to scientific writing, combinatory dictionary of scientific usage*. Russie : Hayka.
- Rinck, F. (2006). *L'article de recherche en Sciences du langage et en Lettres [Texte imprimé] : figure de l'auteur et identité disciplinaire du genre* (Thèse de doctorat). Université Stendhal Grenoble 3, Grenoble 3.
- Rinck, F. (2010). L'analyse linguistique des enjeux de connaissance dans le discours scientifique. *Revue d'anthropologie des connaissances*, 4(3), 427-453.
- Rossari, C. (1990). Projet pour une typologie des opérations de reformulation. *Cahiers de linguistique française*, (11), 345-359.
- Rossari, C. (1992). De l'exploitation de quelques connecteurs reformulatifs dans la gestion des articulations discursives. *Pratiques*, (25), 111-125.
- Rossari, C. (1994). *Les opérations de reformulation : analyse du processus et des marques dans une perspective contrastive français-italien* (2e éd.). Berne ; Berlin : P. Lang.
- Rossari, C., & Jayez, J. (1997). Connecteurs de conséquence et portée sémantique. *Cahiers de linguistique française*, (19), 233-266.
- Roulet, E. (1987). Complétude interactive et connecteurs reformulatifs. *Cahiers de linguistique française*, (8), 111-140.
- Salem, A. (1987). *Pratique des segments répétés : essai de statistique textuelle*. Paris : Klincksieck.
- Sarda, L., & Charolles, M. (2012). Les adverbiaux prépositionnels : position, fonction et portée. Présentation du numéro. *Travaux de linguistique*, (64), 7-20.
- Siepmann, D. (2005). *Discourse markers across languages : a contrastive study of second-level discourse markers in native and non-native text with implications for general and pedagogic lexicography*. London ; New York : Routledge.
- Siepmann, D. (2007). Les marqueurs de discours polylexicaux en français. *Revue française de linguistique appliquée*, XII, 123-136.
- Simard, C. (1994). Pour un enseignement plus systématique du lexique, (92), 28-33.
- Simpson-Vlach, R., & Ellis, N. C. (2010). An Academic Formulas List: New Methods in Phraseology Research. *Applied Linguistics*, 31(4), 487-512.
- Sinclair, J. M. (1991). *Corpus, Concordance, Collocation*. Oxford : Oxford University Press.
- Sinclair, J. M. (2005). Corpus and text : basic principles. In M. Wynne (Éd.), *Developping linguistic corpora : A guide for good practice*. Oxford : Oxbow Books, 1-16.
- Steuckardt, A., & Niklas-Salminen, A. (2005). *Les marqueurs de glose*. Aix-en-Provence, France: Publications de l'Université de Provence.
- Stubbs, M. (2002). Two quantitative methods of studying phraseology in English. *International Journal of Corpus Linguistics*, 7(2), 215-244.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge, United Kingdom : Cambridge University Press.

- Teufel, S. (2000). *Argumentative Zoning: Extraction information from scientific text* (Thèse de doctorat). University of Edinburgh, Edinburgh.
- Thoiron, P., & Béjoint, H. (1991). La place des reformulations dans les textes scientifiques. *Méta : journal des traducteurs*, 36(1), 101-110.
- Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam : Benjamins.
- Tran, T. T. H. (2010). *Aides logicielles à la rédaction universitaire. L'exemple des collocations* (Mémoire Master 2 Recherche FLE). Université Stendhal - Grenoble 3, Grenoble.
- Tutin, A. (2014). La phraséologie transdisciplinaire des écrits scientifiques : des collocations aux routines sémantico-rhétoriques. In A. Tutin & F. Grossmann (Éd.), *L'écrit scientifique : du lexique au discours. Autour de scientext*. Rennes : Presses Universitaires de Rennes, 27-44.
- Tutin, A. (2010). *Sens et combinatoire lexicale: de la langue au discours* (Thèse d'habilitation à diriger des recherches). Université Stendhal Grenoble 3, Grenoble.
- Tutin, A. (2008). Sémantique lexicale et corpus : l'étude du lexique transdisciplinaire des écrits scientifiques. *Lublin Studies in Modern Languages and Literature*, (32), 242-260.
- Tutin, A. (2007a). Modélisation linguistique et annotation des collocations : une application au lexique transdisciplinaire des écrits scientifiques. In S. Koeva, D. Maurel, & M. Silberztein (Éd.), *Formaliser les langues avec l'ordinateur : de Intex à NooJ*. Besançon : Presses Universitaires de Franche Comté, 189-215.
- Tutin, A. (2007b). Traitement sémantique par analyse distributionnelle des noms transdisciplinaires des écrits scientifiques. Présenté à *Actes de TALN*, Toulouse, 5-8.
- Tutin, A., & Falaise, A. (2013). Multiword expressions in scientific discourse : a corpus-driven database. In *Actes de eLex*, Tallinn, Estonie, Tallinn, Estonie, 2013.
- Tutin, A., & Grossmann, F. (2002). Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif. *Revue française de linguistique appliquée*, VII(1), 7-25.
- Vassiliadou, H. (2008). Quand les voies de la reformulation se croisent pour mieux se séparer : à savoir, autrement dit, c'es-à-dire, en d'autres termes. In M.-C. Le Bot, M. Shuwer, & E. Richard (Éd.), *La reformulation : marqueurs linguistiques, stratégies énonciatives*. Rennes : Presses universitaires de Rennes, 35-49.
- Verlinde, S. (2013). *Interactive LanguageToolbox*. Démonstration présenté à *Premières Rencontres phraséologiques, Des collocations aux séquences figées*, Grenoble, 13-15 novembre 2013.
- Virtanen, T. (2004). Point of departure: Cognitive aspects of sentence-initial adverbials. In Virtanen, T. (Éd.), *Approaches to Cognition through Text and Discourse*. Mouton de Gruyter : Berlin/New York, 79-97.
- Wray, A. (2002). *Formulaic sequences and the lexicon*. Cambridge : Cambridge University Press.

Site internet

AERES : <http://www.aeres-evaluation.fr/index.php/Publications/Methodologie-de-l-evaluation/Listes-de-revues-SHS-sciences-humaines-et-sociales>.

Cairn : <http://www.cairn.info/>

Emolex (site du projet) : www.emolex.eu/

Emobase : <http://emolex.u-grenoble3.fr/emoBase/>

Interactive Language Toolbox : <https://ilt.kuleuven.be/inlato/>

Scientext : <http://scientext.msh-alpes.fr/scientext14/>

Scribplus : <http://scribplus.ovh.org/>

Trésor de la langue française informatisé TLFi : <http://www.atilf.fr/tlfi.htm>

Index

A

annotation, 58, 59, 60, 61, 64, 280
apprentissage sur corpus, 3, 211, 212, 214, 217, 274
 approche comparée, 75

B

bottom-up
 stratégies de lecture, 227, 233, 236, 237, 238, 239,
 251, 252, 260, 265, 270

C

cadres organisationnels, 46
 cadres qualitatifs, 46
 cadres thématiques, 46
cohesive devices, 23
 concordancier, 3, 67, 223, 231, 241, 242, 243, 244, 245,
 250, 251, 258
connecteurs logiques, 28, 34, 38, 40, 43

D

data driven-learning, 3
data-based learning, 211
data-driven learning, 211, 282, 285, 287
 déductive
 approche, 218, 219, 227, 233, 234, 235, 236, 239,
 244, 245, 246, 252, 257, 259, 260, 261, 265, 274,
 331
 discours scientifique, 8, 9, 13, 21, 94, 97
 Discursivité, 73

E

écrits scientifiques, 2, 3, 4, 5, 7, 8, 9, 10, 11, 13, 14, 17,
 19, 20, 21, 22, 25, 27, 30, 34, 39, 40, 43, 50, 53, 54,
 55, 56, 58, 77, 78, 79, 81, 89, 93, 98, 99, 100, 105,
 203, 207, 214, 215, 217, 218, 220, 222, 223, 224, 225,
 228, 236, 237, 239, 262
 écrits universitaires, 1, 5, 7, 10, 11, 40, 87, 217
 empirique, 5, 19, 56, 75, 76, 118, 124

F

fonctions rhétoriques, 14, 19, 21, 218, 224, 225, 228, 262,
 265
 Français sur Objectif Spécifique, 33

I

idiomes, 27
 inductive
 approche, 211, 218, 219, 220, 221, 223, 224, 227,
 228, 233, 234, 235, 236, 239, 243, 244, 250, 251,
 252, 257, 258, 259, 261, 265, 266, 270, 271, 274,
 331

L

lexical bundles, 23, 63
 linguistique de corpus, 2, 4, 53, 109, 214, 229

M

marqueurs d'addition, 81, 82, 177
 marqueurs d'énumération, 48, 81, 83, 186, 280
 Marqueurs d'exemplification, 87
 marqueurs d'opposition, 91
 marqueurs de causalité, 90, 265
 Marqueurs de comparaison, 89, 257
 marqueurs de concession, 91, 261, 270
 marqueurs de conclusion, 85
 marqueurs de condition, 92
 marqueurs de finalité, 92
 marqueurs de la reformulation non paraphrastique, 101,
 104
 marqueurs de reformulation, 4, 5, 6, 35, 36, 38, 77, 85,
 87, 93, 94, 96, 98, 99, 103, 105, 106, 107, 108, 111,
 116, 117, 119, 120, 121, 123, 200, 202, 207, 208, 209,
 221, 225, 234, 241, 257, 259, 262, 274, 276, 277, 279
 marqueurs de reformulation non paraphrastiques, 103
 marqueurs de reformulation paraphrastiques, 100
 marqueurs de structuration, 45, 83
 marqueurs de topicalisation, 45, 81, 85, 93, 221, 249,
 250, 251, 252, 253, 260, 261, 262, 267, 270, 271, 272,
 280
 marqueurs polylexicaux, 48, 206, 207, 279
 métadiscours, 10, 19

O

onomasiologique
 approche, 16, 21, 22, 40, 41, 42, 50, 51, 74, 76, 77, 79,
 107, 218, 219, 220, 221, 223, 224, 225, 227, 230,
 231, 233, 235, 239, 243, 244, 253, 261, 266, 270,
 271, 274, 276

S

schémas collocationnels, 17, 18
segments répétés, 15, 20, 61, 63, 64, 68
sémasiologique, 21, 40, 42, 50, 74, 218, 225, 227, 230,
231, 233, 234, 235, 236, 239, 243, 245, 246, 252, 265,
270, 274
stratégies d'enseignement, 32, 202, 211, 214, 215, 218,
227
stratégies de lecture, 177, 200, 224, 226, 233, 237, 239,
247, 249, 251
syntactico-sémantique, 16, 17, 76, 107, 108, 117, 119,
155, 202, 244, 250, 259, 262

T

top-down
stratégies de lecture, 227, 228, 236, 237, 238, 239,
250, 251, 252, 261, 265, 271
transdisciplinaire, 1, 2, 4, 5, 7, 11, 13, 14, 20, 21, 22, 25,
27, 54, 58, 61, 70, 71, 72, 77, 78

U

unités polylexicales, 3, 13, 17, 20, 24, 25, 39, 48, 73, 80
univers de discours, 46

Annexe

Annexe 1: Typologie des séquences lexicalisées à fonction discursive

- Cette typologie recense au total 171 séquences lexicalisées à fonction discursive
- NB : Pour les catégories syntaxiques :
 - L.P. : *Locution prépositive*
 - L.A. : *Locution adverbiale*
 - L.C. : *Locution conjonctive*
 - S.P. : *Structure phrastique*

TYPES DE MARQUEURS		Séquences lexicalisées à fonction discursive	Catégories syntaxiques	Fréquence	Discipline
FONCTION METADISCURSIVE					
Enchaînement thématique	Addition	Ces marqueurs servent à signaler un changement de perspective ou non sur le thème développé (8 MD).			
		ainsi que	L.C.	900	10
		par ailleurs	L.A.	498	10
		de plus	L.A.	291	10
		en outre	L.A.	277	10
		entre autres	L.A.	89	10
		de surcroît	L.A.	57	9
		en plus	L.A.	48	9
		au demeurant	L.A.	31	9
	Structuration	Ces marqueurs servent à mettre en relief deux aspects (complémentaires ou opposés) d'un élément (4 MD).			
		d'une part	L.A.	456	10
		d'autre part	L.A.	277	10
		d'un côté	L.A.	133	10

		d'un autre côté	L.A.	41	9
	Topicalisation	Ce type de marqueur permet à l'auteur d'indiquer un changement de point de vue, donc le passage d'une unité de discours à une autre (12 MD).			
		en termes de	L.P.	363	10
		à propos de + N	L.P.	305	10
		à ce propos	L.A.	302	10
		en matière de + N	L.P.	269	10
		en ce qui concerne + N	L.P.	201	10
		au sujet de + N	L.P.	84	8
		à l'exception de	L.P.	70	10
		à cet égard	L.A.	49	10
		à ce sujet	L.A.	41	10
		il est question de + N	S.P.	36	8
		sur le plan	L.P.	33	10
		en guise de	L.P.	19	6
	Énumération	Ces marqueurs permettent d'énoncer successivement les parties d'un tout, le segment textuel est donc considéré comme l'un des termes d'une énumération (14 MD).			
		tout d'abord	L.A.	217	10
		pour commencer	L.A.	14	5
		par la suite	L.A.	127	10
		en second lieu	L.A.	22	9
		en premier lieu	L.A.	88	10
		dans un second temps	L.A.	49	9
		dans un premier temps	L.A.	138	10
		dans la suite de + N	L.P.	37	10
		d'abord	L.A.	144	10
		d'emblée	L.A.	85	10
		avant tout	L.A.	172	10
		au premier abord	L.A.	11	7
		au préalable	L.A.	17	10
		à commencer par + N	L.P.	32	7

	Conclusion	Ces marqueurs servent à marquer le dernier élément d’une série (4 MD).			
		En conclusion		24	8
		Pour conclure		24	9
		Pour finir		10	6
Reformulation	Le marqueur de reformulation permet à l’auteur d’élucider le contenu par reformulation. La reformulation fournit d’amples informations afin d’apporter au lecteur une meilleure compréhension du message de l’auteur. (12 MD)				
	c’est-à-dire	L.A.	704	10	
	à savoir	L.A.	141	10	
	autrement dit	L.A.	197	10	
	en d’autres termes	L.A.	119	10	
	En somme	L.A.	57	9	
	En résumé	L.A.	22	9	
	En un mot	L.A.	15	7	
	Au total	L.A.	59	9	
	En définitive	L.A.	54	10	
	Au fond	L.A.	15	5	
	Somme toute	L.A.	13	7	
	En fin de compte	L.A.	11	6	
Exemplification	Le marqueur d’exemplification consistent en des propositions ou des groupes prépositionnels qui introduisent des exemples afin d’illustrer, prouver, éclairer une idée, un sujet. (7 MD)				
	par exemple	L.A.	1357	10	
	tel(le) que	L.C.	609	10	
	en particulier	L.A.	582	10	
	à titre d’exemple	L.A.	27	10	
	pour illustrer	L.A.	24	10	
	toutes choses égales par ailleurs	L.A.	22	5	
	à titre d’illustration	L.A.	10	6	
	à l’exemple de + N	L.P.	10	4	

Comparaison		Le marqueur de comparaison sert à examiner deux ou plusieurs sujets, entre le comparant et le comparé pour en établir les ressemblances et les différences. (16 MD)			
		par rapport à + N	L.P.	626	10
		d'autant que + pp V à l'indicatif	L.C.	343	10
		d'autant plus... que + pp	L.C.	279	10
		autant que + N	L.C.	114	10
		à l'instar de + N	L.P.	107	9
		de même que + pp V à l'indicatif	L.C.	101	10
		au même titre que + N	L.C.	64	10
		conformément à + N	L.P.	64	10
		il en va de même que + pp V à l'indicatif	S.P.	34	9
		en différence de + N	L.A.	32	4
		il en est de même + N	S.P.	28	8
		toutes choses égales par ailleurs	L.A.	22	5
FONCTION ARGUMENTATIVE					
Causalité	Cause	Le marqueur de causalité sert à instaurer le lien de cause à effet. Les groupes adverbiaux signalent la raison ou le motif du sujet abordé. (15 MD)			
		en effet	L.A.	1493	10
		parce que + pp V à l'indicatif	L.C.	743	10
		d'autant plus... que + pp V à l'indicatif	L.C.	279	10
		dans la mesure où + pp V à l'indicatif	L.C.	255	10
		en raison de + N	L.P.	195	10
		du fait de + N	L.P.	182	10

		compte tenu de + N	L.P.	106	10
		par le fait que + N	L.C.	70	10
		pour cette raison	L.A.	42	10
		ce qui explique + pp V à l'indicatif	S.P.	37	9
		ce qui signifie que + pp V à l'indicatif	S.P.	35	10
		vu que + pp V à l'infinitif	L.C.	32	9
		en vertu de + N	L.P.	28	9
		étant donné que + pp V à l'indicatif	L.C.	23	9
		d'autant mieux... que + pp V à l'indicatif	L.C.	17	5
	Conséquence	Le marqueur de conséquence exprime une suite logique entraînée par un fait. (12 MD)			
		d'où	L.P.	210	10
		par conséquent	L.A.	165	10
		c'est pourquoi	L.C.	116	10
		de ce fait	L.A.	99	10
		en conséquence	L.A.	86	10
		de façon à ce que + pp V au subjonctif	L.P.	67	10
		de sorte que + pp V à l'indicatif	L.C.	65	8
		au point de + V au point que + pp V à l'indicatif	L.P.	59	9
		si bien que + pp V à l'infinitif	L.C.	50	10
		il en résulte + N /il en résulter que + pp V au subjonctif	S.P.	39	8
		à tel point que + pp V à	L.C.	7	5

		l'infinitif			
		de telle façon que + pp V au subjonctif	L.C.	6	4
Concession	Le marqueur de concession sert à signaler une opposition, une restriction à l'action exprimée par la principale. (36 MD)				
	d'ailleurs	L.A.	545	10	
	par ailleurs	L.A.	495	10	
	même si + pp V à l'indicatif	L.A.	424	10	
	en revanche	L.A.	403	10	
	bien que + pp V au subjonctif	L.C.	388	10	
	de fait	L.A.	228	10	
	en fait	L.A.	125	10	
	sans que + pp V au subjonctif	L.C.	124	10	
	en dépit de + N	L.P.	90	9	
	au lieu de + V à l'infinitif	L.P.	84	10	
	en quelque sorte	L.A.	84	10	
	en tout cas	L.A.	72	10	
	alors même que + pp V à l'indicatif	L.C.	60	10	
	quoique	L.C.	48	8	
	cependant que + pp V à l'indicatif	L.C.	47	10	
	quand même	L.A.	44	8	
	malgré tout	L.A.	44	10	
	tout de même	L.A.	33	6	
	à moins de + V à l'infinitif	L.P.	31	9	
	dans une certaine mesure	L.A.	29	8	
	dans une certaine mesure	L.A.	29	8	
	dans une moindre mesure	L.A.	28	8	
	dans une moindre mesure	L.A.	28	8	

	en aucun cas	L.A.	24	9
	d'une certaine manière	L.A.	24	9
	en aucun cas	L.A.	24	9
	d'une certaine manière	L.A.	24	9
	quand bien même + pp V au conditionnel	L.A.	20	7
	quand bien même	L.A.	20	7
	pour peu que + pp V au subjonctif	L.C.	19	7
	d'une certaine façon	L.A.	19	7
	en tout état de cause	L.A.	16	8
	à moins que + V au subjonctif	L.C.	15	8
	à moins que + pp V au subjonctif	L.C.	15	8
	pour le moins	L.A.	14	7
	à tout le moins	L.A.	14	6
Opposition	Le marqueur d'opposition sert à exprimer une opposition simple entre les deux fait. (16 MD)			
	alors que + pp V à l'infinitif	L.C.	938	10
	au contraire de + N	L.P.	726	10
	tandis que + V à l'indicatif	L.C.	349	10
	contrairement à + N	L.P.	200	10
	à l'inverse	L.A.	157	10
	pour autant sans pour autant	L.P.	107	10
	par contre	L.A.	86	10
	à la différence de + N	L.P.	72	10
	à l'encontre de + N	L.P.	69	9
	par opposition / par opposition à + N	L.A. / L.P.	60	10

	par contraste / par le contraste avec + N	L.A. / L.P.	29	8
	en contradiction / en contradiction avec + N	L.P.	21	8
	a contrario	L.A.	17	8
	à l'opposé / à l'opposé de + N	L.P.	15	6
Condition	Le marqueur de condition sert à indiquer la ou les conditions pour que l'action se produise. (8 MD)			
	dans le cas où +	L.P.	264	10
	en cas de + N	L.P.	113	10
	pour autant que + pp V au subjonctif	L.C.	30	10
	à la condition de + N	L.P.	17	6
	à supposer que + pp V au subjonctif	L.C.	17	6
	à la condition que + pp V au subjonctif	L.C.	16	7
	dans l'hypothèse où + pp V au conditionnel	L.C.	8	5
	sous réserve de + N	L.P.	7	4
Finalité	L'objectif de l'action est véhiculé soit par des groupes adverbiaux, prépositionnels soit par des groupes verbaux. (10 MD)			
	afin de + V à l'infinitif	L.P.	581	10
	pour que + pp V au subjonctif	L.C.	255	10
	en vue de + N	L.P.	105	10
	au moyen de + V à l'infinitif	L.P.	81	10
	de façon à + V à l'infinitif	L.P.	67	10
	de manière à + V à	L.P.	62	10

	l'infinitif			
	dans le but de + V à l'infinitif	L.P.	61	10
	à cet effet	L.A.	28	10
	en sorte que + pp V au subjonctif	L.C.	20	9
	dans le dessein de + V à l'infinitif	L.P.	11	3

Annexe 2 : Liste des revues en SHS

Annexe 1.1 : Linguistique

Cussac Pomel, J., & Garcia Bourgade, G. (2006). Les procédés d'autocopie de l'enfant apprenti scripteur. *Autour des langues et du langage: perspective pluridisciplinaire*. Colloque des Etudiants chercheurs en Didactique des Langues et en Linguistique, 4-7 juillet 2007. Consulté le 20 octobre 2014 : <http://w3.erss.univ-tlse2.fr/textes/publications/CDG/25/CG25-8-Bourigault.pdf>. [lin-art-530]

Lecolle, M. (2000). Figures et références plurielle en corpus journalistique. *Cahiers de Grammaire*, 25, 29-52. [lin-art-269]

Bourigault, D., & Fabre, C. (2000). Approche linguistique pour l'analyse syntaxique de corpus. *Cahiers de Grammaire*, 25, 131-151. Consulté le 20 octobre 2014 : <http://w3.erss.univ-tlse2.fr/textes/publications/CDG/25/CG25-8-Bourigault.pdf>. [lin-art-75]

Aussenac-Gilles, N., & Séguéla, P. (2000). Les relations sémantiques : du linguistique au formel. *Cahiers de Grammaire*, 25, 175-198. Consulté le 20 octobre 2014 : <http://w3.erss.univ-tlse2.fr/textes/publications/CDG/25/CG25-10-Aussenac.pdf>. [lin-art-50]

Rebeyrolle, J., & Tanguy, L. (2000). Repérage automatique de structures linguistiques en corpus : le cas des énoncés définitoires. *Cahiers de Grammaire*, 25, 153-174. Consulté le 20 octobre 2014 : <http://w3.erss.univ-tlse2.fr/textes/publications/CDG/25/CG25-9-Rebeyrolle.pdf>. [lin-art-169]

Niederberger Costello, N., & Berthoud-Papandropoulou, I. (2004). Utilisation des pronoms personnels en français écrit par des enfants sourds bilingues : un parcours spécifique d'apprentissage ?. *Acquisition et enseignement de la morphographie, LIDIL*, 30, 27-38. [lin-art-488]

Pellat, J-C., & Teste, G. (2004). Morphographie et production d'écrits au cycle 3 des écoles. *Acquisition et enseignement de la morphographie, LIDIL*, 30, 87-100. [lin-art-235]

Payre-Ficout, C., & Chevrot, J-P. (2004). La forme contre l'usage: étude exploratoire de l'acquisition du prétérit anglais. *Acquisition et enseignement de la morphographie, LIDIL*, 30, 101-116. [lin-art-225]

Biichlé, L. (2006). Représentations de migrants originaires du Maghreb : distinction des variétés diglossiques par les représentations sur la production et la réception. *Autour des langues et du langage: perspective pluridisciplinaire*. Colloque des Etudiants chercheurs en Didactique des Langues et en Linguistique, 4-7 juillet 2007. [lin-art-159]

Delanoë, A. (2006). Les troubles pragmatiques et émotionnels chez des enfants et adolescents cérébrolésés : l'exemple de la situation de production narrative. *Autour des langues et du langage: perspective pluridisciplinaire*. Colloque des Etudiants chercheurs en Didactique des Langues et en Linguistique, 4-7 juillet 2007. [lin-art-105]

Gautier, A. (2006). Le concept de phrase canonique en linguistique française. *Autour des langues et du langage: perspective pluridisciplinaire*. Colloque des Etudiants chercheurs en Didactique des Langues et en Linguistique, 4-7 juillet 2007. [lin-art-424]

Blanche-Benveniste, C. (2001). Préposition à éclipses. *Travaux de Linguistique*, 42-43, 83-96. Consulté le 20 octobre 2014 : <http://www.cairn.info/article.php?REVUE=travaux-de-linguistique&ANNEE=2001&NUMERO=1&PP=83>. [lin-art-305]

Kupferman, L. (2001). Les déplacements des syntagmes en de: un regard du troisième type. *Travaux de Linguistique*, 42-43, 33-41. Consulté le 20 octobre 2014 : <http://www.cairn.info/revue-travaux-de-linguistique-2001-1-page-33.htm>. [lin-art-67]

Anscombre, J-C. (2001). L'analyse de la construction En tout N par D. Leeman : Quelques remarques. *Travaux de Linguistique*, 42-43, 183-198. Consulté le 20 octobre 2014 : <http://www.cairn.info/revue-travaux-de-linguistique-2001-1-page-183.htm>. [lin-art-369]

- Bacha, J. (2000). Marqueurs exclamationnels et aspect verbal. *Revue de Sémantique et Pragmatique*, 7, 9-27. [lin-art-621]
- Bergounioux, G. (2002). Endophasie et linguistique. Décomptes, quotes et squelette. *Langue Française*, 132-1, 106-124. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/lfr_0023-8368_2001_num_132_1_6318. [lin-art-508]
- Cusin-Berche, F. (1997). Sens et contre sens : du cotexte au contexte. *Revue de Sémantique et Pragmatique*, 1, 135-146. [lin-art-345]
- De Vogüé, Franckel (2002). Identité et variation de l'adjectif grand. *Langue Française*, 133-1, 28-41. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/lfr_0023-8368_2002_num_133_1_1044. [lin-art-528]
- Dominicy, M., & Martin, F. (2001). À travers, au travers (de) et le point de vue. *Travaux de Linguistique*, 42-43, 211-227. Consulté le 20 octobre 2014 : http://www.cairn.info/resume.php?ID_ARTICLE=TL_042_221. [lin-art-136]
- Franckel, J.-J., & Paillard, D. (1997). Représentation formelle des mots du discours. Le cas de D'ailleurs. *Revue de Sémantique et Pragmatique*, 1, 51-64. [lin-art-176]
- Godart-Wendling, B. (2000). Comment ça réfère ?. *Revue de Sémantique et Pragmatique*, 7, 105-121. [lin-art-306]
- Jalenques, P. (2002). Etude sémantique du préfixe RE en français contemporain : A propos de plusieurs débats actuels en morphologie dérivationnelle. *Langue Française*, 133-1, 74-90. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/lfr_0023-8368_2002_num_133_1_1048. [lin-art-490]
- Kleiber, G. (1997). Contexte, où es-tu ?. *Revue de Sémantique et Pragmatique*, 1, 65-79. [lin-art-411]
- Léon, J. (2003). Approche séquentielle d'un objet sémantico-pragmatique : le couple Q-R, questions alternatives et questions rhétoriques. *Revue de Sémantique et Pragmatique*, 1, 23-50. Consulté le 20 octobre 2014 : <http://htl.linguist.univ-paris-diderot.fr/leon/Leon1997b.pdf>. [lin-art-38]
- Malrieu, D. (1997). Reconnaissance de catégories discursives et repérages énonciatifs : l'exemple des hypothèses dans les textes scientifiques. *Revue de Sémantique et Pragmatique*, 1, 147-163. [lin-art-622]
- Miche, E. (2000). L'articulation entre les structures linguistique, textuelle et situationnelle du discours. *Revue de Sémantique et Pragmatique*, 7, 165-178. [lin-art-99]
- Moeschler, J. (2000). Pragmatique. Etat de l'art et perspectives. *Marges linguistiques*, 1, <http://www.marges-linguistiques.com>. Consulté le 20 octobre 2014 : http://www.unige.ch/lettres/linguistique/moeschler/publication_pdf/pragmatique.pdf. [lin-art-509]
- Porhiel, S. (1997). Au sujet de et à propos de - une analyse lexicographique, discursive et linguistique. *Travaux de Linguistique*, 42-43, 171-181. Consulté le 20 octobre 2014 : <http://www.cairn.info/revue-travaux-de-linguistique-2001-1-page-171.htm>. [lin-art-68]
- Puech, C. (2002). Langage intérieur et ontologie linguistique à la fin du XIXe siècle. *Langue Française*, 132-1, 26-47. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/lfr_0023-8368_2001_num_132_1_6313. [lin-art-42]
- Romero-Lopez, M.-C. (2002). Identité et variation du verbe jouer. *Langue Française*, 133-1, 63-73. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/lfr_0023-8368_2002_num_133_1_1047. [lin-art-431]

Annexe 1.2 : Psychologie

- Sanguin-Bruckert, C., & Bruckert, J.-P. (2004). Le rôle des connaissances morphographiques. *Acquisition Et Enseignement De La Morphographie, LIDIL*, 30, p147-168. Consulté le 20 octobre 2014 : <http://lidil.revues.org/index893.html>. [psy-art-409]

- Auvray, M., & O'regan, J-K. (2003). L'influence des facteurs sémantiques sur la cécité aux changements progressifs dans les scènes visuelles. *L'année psychologique*, Vol. 103, 1, 9-32. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2003_num_103_1_29621. [psy-art-185]
- Bocéréan, C. (2001). La relation de similarité à la base du choix des distracteurs : contribution à la généralisation du modèle de rumelhart et abrahamson. *L'année psychologique*, Vol. 101, 1, 33-63. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2001_num_101_1_29715. [psy-art-606]
- Butera, F., & Darnon, B. (2005). Buts d'accomplissement, stratégies d'étude et motivation intrinsèque: présentation d'un domaine de recherche et validation française de l'échelle d'elliot mcgregor (2001). *L'année psychologique*, Vol. 105, 1, 105-131. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2005_num_105_1_3821. [psy-art-449]
- Catteau, V., & Chabrol, H. (2005). Etude des relations entre les stratégies d'adaptation aux sentiments dépressifs, la symptomatologie dépressive et les idées suicidaires chez l'adolescent. *L'année psychologique*, Vol 105, 3, 451-476. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2005_num_105_3_29704. [psy-art-178]
- Ciccotti, S. (2004). Les orateurs ont-ils intérêt à s'épiler les sourcils? influence d'indices non verbaux sur les processus persuasifs. *L'année psychologique*, Vol 104, 2, 227-247. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2004_num_104_2_29665. [psy-art-260]
- Courbois, Y. (2005). Le développement de la " rotation mentale " effets de la saillance de l'axe du stimulus et de son alignement avec la direction verticale. *L'année psychologique*, Vol 105, 1, 369-386. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2005_num_105_1_3819. [psy-art-534]
- Dalle, N., & Niedenthal, P-M. (2003). La réorganisation de l'espace conceptuel au cours des états émotionnels. *L'année psychologique*, Vol 103, 4, 585-616. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2003_num_103_4_29653. [psy-art-246]
- Detable, C., & Vinter, A. (2004). Le maintien dans le temps des effets d'un apprentissage implicite chez des enfants et adolescents avec retard mental. *L'année psychologique*, Vol 104, 4, 751-770. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2004_num_104_4_29687. [psy-art-53]
- Facon, R. (2003). Sur la loi des rendements décroissants. Efficience intellectuelle et facteur général. *L'année psychologique*, Vol 103, 1, 81-102. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2003_num_103_1_29624. [psy-art-296]
- Font, N., & Lavaur, J-M. (2004). Effets de la fréquence du voisinage orthographique interlangue lors de la reconnaissance visuelle de mots chez les bilingues. *L'année psychologique*, Vol 104, 3, 377-405. [psy-art-539]
- Giraudet, G., & Roumes, C. (2004). La signature spatiale de l'objet une information essentielle pour la localisation de cibles dans une scène naturelle. *L'année psychologique*, Vol 104, 1, 9-49. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2004_num_104_1_3926. [psy-art-442]
- Martinie, M-E., & Joule, J-V. (2004). Changement d'attitude et fausse attribution effet de la centration sur le comportement de soumission. *L'année psychologique*, Vol 104, 3, 517-535. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2004_num_104_3_29677. [psy-art-428]
- Jover, M., & Mellier, D. (2005). Connaissance du contexte et anticipation posturale chez l'enfant au cours du développement. *L'année psychologique*, Vol 105, 4, 553-572. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2005_num_105_4_30492. [psy-art-579]
- Largy, P. (2001). La révision des accords nominal et verbal chez l'enfant. *L'année psychologique*, Vol 101, 2, 221-245. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2001_num_101_2_29554. [psy-art-501]

- Lemetayer, F. (2005). Le développement de l'autorégulation du comportement dans un contexte interactionnel au cours de la deuxième année de vie. *L'année psychologique*, Vol 105, 4, 573-590. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2005_num_105_4_30493. [psy-art-382]
- Montfort, V., & Pouthas, V. (2005). Interruption dans la reproduction d'une durée : effet sur la précision du jugement temporel. *L'année psychologique*, Vol 105, 2, 209-224. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2005_num_105_2_29692. [psy-art-13]
- Rabillard, J-F., & Corson, Y. (2005). Ressemblance et catégorisation: étude de l'effet d'un contexte de nature axiologique. *L'année psychologique*, Vol 105, 4, 591-624. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2005_num_105_4_30494. [psy-art-366]
- Redersdorff, S., & Martinot, S. (2003). Impact des comparaisons ascendantes et descendantes sur l'estime de soi importance de l'identité mise en jeu. *L'année psychologique*, Vol 103, 3, 411-443. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2003_num_103_3_29643. [psy-art-194]
- Lhote, L., & Streri, A. (2003). La mémoire haptique de la forme des objets chez les bébés âgés de 4 mois. *L'année psychologique*, Vol 103, 1, 33-49. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2003_num_103_1_29622. [psy-art-113]
- Vieillard, S., & Bougeant, J-C. (2005). Performances à une tâche de mémoire de travail sous induction émotionnelle négative influence modulatrice de l'état émotionnel sur les processus exécutifs. *L'année psychologique*, Vol 105, 1, 63-104. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/psy_0003-5033_2005_num_105_1_3820. [psy-art-407]
- Doignon, N., & Zagar, D. (2006). Les enfants en cours d'apprentissage de la lecture perçoivent-ils la syllabe a l'écrit ?. *Canadian journal of experimental psychology*, Vol. 60, 4, 258-274. [psy-art-303]
- Reilhac, G., & Jiménez, M. (2006). Amorçage de la couleur typique d'un objet lors d'une tâche de catégorisation. *Canadian journal of experimental psychology*, Vol. 60, 4, p285-296. [psy-art-371]
- Brechet, C., Picard, D., & Baldy, R. (2007). Expression des émotions dans le dessin d'un homme chez l'enfant de 5 à 11 ans. *Canadian journal of experimental psychology*, Vol 61, 2, 142-155. [psy-art-492]
- Villieux, A., & Delhomme, P. (2009). Colère éprouvée au volant et différentes manières de l'exprimer : quels liens avec des transgressions de conduite déclarées ?. *Revue Le travail Humain*, Vol 71, 4, 359-384. Consulté le 20 octobre 2014 : http://www.cairn.info/resume.php?ID_ARTICLE=TH_714_0359. [psy-art-318]
- Fromage, Louaisil (2001). De l'espace à l'espace habité ou l'histoire d'un déménagement. *Revue Québécoise de Psychologie*, Vol 2, 3. [psy-art-583]

Annexe 1.3 : Sciences de l'Éducation

- Talbot, L. (2004). Intérêts et limites des apports de la recherche aux pratiques de l'enseignement. *Revue recherches et éducations*, 8. Consulté le 20 octobre 2014 : <http://rechercheseducations.revues.org/index354.html>. [sed-art-104]
- Biemar, S. (2008). Etude des images identitaires de la relation pédagogique lors de l'insertion professionnelle. *Revue Education et Formation en ligne*, Vol e-289, 57-30. [sed-art-81]
- Archer, C. (2008). L'enseignement supérieur: un double filtrage social?. *Revue Education et Formation en ligne*, Vol. e-287, 43-53. [sed-art-229]
- Perez-Roux, T. (2008). Evolution des représentations et des pratiques dans la formation initiale des enseignants: éléments de construction d'une identité professionnelle à l'échelle du temps. *Revue Education et Formation en ligne*, Vol. e-287, 7-19. [sed-art-362]

- Alexander, R. (2003). Pédagogie, culture et comparaison : visions et versions de l'école élémentaire. *Revue Française de Pédagogie*, 142, 5-19. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_142_1_2927. [sed-art-417]
- Troger, V. (2003). Des identités professionnelles à l'épreuve du changement : les formateurs de l'enseignement technique et professionnel en IUFM. *Revue Française de Pédagogie*, 142, 79-88. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_142_1_2934. [sed-art-412]
- Meuret, D. (2005). Pourquoi les jeunes Français ont-ils à quinze ans des compétences inférieures à celles de jeunes d'autres pays ? *Revue Française de Pédagogie*, 142, 89-104. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_142_1_2936. [sed-art-467]
- Demailly, L. (2003). L'évaluation de l'action éducative comme apprentissage et négociation. *Revue Française de Pédagogie*, 142, 115-129. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_142_1_2938. [sed-art-234]
- Mons, N. (2004). Politiques de décentralisation en éducation : diversité internationale, légitimations théoriques et justifications empiriques. *Revue Française de Pédagogie*, 146, 41-52. Consulté le 20 octobre 2014 : http://ife.ens-lyon.fr/publications/edition-electronique/revue-francaise-de-pedagogie/INRP_RF146_4.pdf. [sed-art-414]
- Orivel, F. (2004). Atteindre l'éducation pour tous en 2015 est-il un objectif réalisable?. *Revue Française de Pédagogie*, 146, 53-63. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2004_num_146_1_3093. [sed-art-464]
- Mangez, E. (2004). La production de programmes de cours par les agents intermédiaires : transferts de savoirs et relations de pouvoir. *Revue Française de Pédagogie*, 146, 65-77. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2004_num_146_1_3095. [sed-art-330]
- Crahay, M. (2004). Peut-on conclure à propos des effets du redoublement ? *Revue Française de Pédagogie*, 148, 11-23. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2004_num_148_1_3246. [sed-art-482]
- Bressoux, P. (2004). Formalisation et modélisation dans les sciences sociales : une étude de la construction du jugement des enseignants. *Revue Française de Pédagogie*, 148, 61-74. Consulté le 20 octobre 2014 : http://ife.ens-lyon.fr/publications/edition-electronique/revue-francaise-de-pedagogie/INRP_RF148_6.pdf. [sed-art-311]
- Guimond, S. (2004). Évaluer l'impact des pratiques pédagogiques : perspectives de la psychologie sociale expérimentale. *Revue Française de Pédagogie*, 148, 25-36. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2004_num_148_1_3247. [sed-art-122]
- Dugas, E. (2004). Des pratiques sociales aux pratiques scolaires en EPS : des logiques de scolarisation plurielles. *Revue Française de Pédagogie*, 149, 5-17. Consulté le 20 octobre 2014 : http://ife.ens-lyon.fr/publications/edition-electronique/revue-francaise-de-pedagogie/INRP_RF149_1.pdf. [sed-art-52]
- Legardez, A. (2004). Transposition didactique et rapports aux savoirs : l'exemple des enseignements de questions économiques et sociales, socialement vives. *Revue Française de Pédagogie*, 149, 19-27. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2004_num_149_1_3169. [sed-art-49]
- Philippe, J. (2004). La transposition didactique en question: pratiques et traduction. *Revue Française de Pédagogie*, 149, 29-36. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2004_num_149_1_3170. [sed-art-574]
- Legendre, F. (2004). Représentations du métier et de l'école des accédants enseignants issus des immigrations à l'IUFM de Créteil. *Revue Française de Pédagogie*, 149, 65-81. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2004_num_149_1_3174. [sed-art-124]

- Marchive, A. (2003). Ethnographie d'une rentrée en classe de cours préparatoire : comment s'instaurent les règles de la vie scolaire ?. *Revue Française de Pédagogie*, 142, 21-32. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_142_1_2928. [sed-art-8]
- Fabre, M. (2003). L'école peut-elle encore former l'esprit ?. *Revue Française de Pédagogie*, 143, 7-15. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_143_1_2947. [sed-art-257]
- Audureau, J-P. (2003). Assujettissement et subjectivation : réflexions sur l'usage de Foucault en éducation. *Revue Française de Pédagogie*, 143, 17-29. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_143_1_2948. [sed-art-540]
- Pierrot, A. (2003). Sens et rationalité à l'école. Quelques réflexions librement inspirées de Wittgenstein. *Revue Française de Pédagogie*, 143, 31-41. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_143_1_2949. [sed-art-47]
- Hameline, D. (2003). À propos d'éducation rationnelle : comment qualifier l'éducation ?. *Revue Française de Pédagogie*, 143, 43-52. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_143_1_2950. [sed-art-3]
- Avazini, G. (2003). Scientificité, axiologie et argumentation chez les théoriciens de l'Éducation Nouvelle. *Revue Française de Pédagogie*, 143, 53-59. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_143_1_2951. [sed-art-383]
- Houssaye, J. (2003). Les tribulations du Bien et du Vrai en éducation. *Revue Française de Pédagogie*, 143, 69-78. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_2003_num_143_1_2953. [sed-art-334]

Annexe 1.4 : Economie

- Béraud, D. (1998). Croissance et endettement dans un modèle à deux pays. *Annales d'économie et de statistique*, 51, 149-168. Consulté le 20 octobre 2014 : <http://www.jstor.org/discover/10.2307/20076141?uid=3738016&uid=2129&uid=2&uid=70&uid=4&sid=56177598143>. [eco-art-325]
- Berthomieu, C., & Ehrhart, C. (2000). Le néostructuralisme comme fondement d'une stratégie de développement alternative aux recommandations néolibérales. *Économie appliquée*, tome LIII, 4, 61-91. [eco-art-275]
- Chatelain, J-B. (1999). Taux de marge et structure financière. *Annales d'économie et de statistique*, 53, 127-147. Consulté le 20 octobre 2014 : <http://halshs.archives-ouvertes.fr/halshs-00118639>. [eco-art-454]
- Curien, N., & Oubejja, K. (1999). Réseaux multiservices : équilibres et dynamique. *Annales d'économie et de statistique*, 53, 213-228. Consulté le 20 octobre 2014 : <http://www.jstor.org/discover/10.2307/20076172?uid=3738016&uid=2129&uid=2&uid=70&uid=4&sid=56177598143>. [eco-art-536]
- De Morogues, F. (1999). Équilibres monétaires du jeu stratégique de marché dans le modèle à générations imbriquées. *Annales d'économie et de statistique*, 54, 1-21. Consulté le 20 octobre 2014 : <http://annales.ensae.fr/anciens/n54/vol54-01.pdf>. [eco-art-39]
- Doz, C., & Lengart, F. (1999). Analyse factorielle dynamique : test du nombre de facteurs, estimation et application à l'enquête de conjoncture dans l'industrie. *Annales d'économie et de statistique*, 54, 91-127. Consulté le 20 octobre 2014 : <http://annales.ensae.fr/anciens/n54/vol54-05.pdf>. [eco-art-188]
- Favard, P., & Bousquet, A. (2001). Hétérogénéité des agents et la relation pollution-revenu. *Revue économique*, Vol. 52, 6, 1185-1203. [eco-art-102]

- Fougère, D., Goux, D., & Maurin, E. (2001). Formation continue et carrières salariales. *Annales d'économie et de statistique*, 62, 49-69. Consulté le 20 octobre 2014 : <http://Annales.ensae.fr/anciens/n62/vol62-04.pdf>. [eco-art-22]
- Gadreau, M. (2001). Les réseaux dans le système de santé et l'arbitrage efficacité-équité. *Économie Appliquée*, tome LIV, 2, 91-128. [eco-art-254]
- Gastaldo, S., & Ragot, L. (2000). Croissance endogène et pollution : une approche fondée sur le comportement du consommateur. *Annales d'Économie et de Statistique*, 57, 25-47. Consulté le 20 octobre 2014 : <http://www.jstor.org/discover/10.2307/20076213?uid=3738016&uid=2129&uid=2&uid=70&uid=4&sid=56177598143>. [eco-art-553]
- Hurlin, C., & Portier, F. (1999). Taux d'actualisation public, distorsions fiscales et croissance endogène. *Annales d'économie et de statistique*, 54, 173-201. Consulté le 20 octobre 2014 : <http://Annales.ensae.fr/anciens/n54/vol54-08.pdf>. [eco-art-78]
- Jayet, P-A. (2001). Évaluation de politiques de retrait des terres pour la régulation d'un marché agricole. *Annales d'économie et de statistique*, 61, 119-150. Consulté le 20 octobre 2014 : <http://Annales.ensae.fr/anciens/n61/vol61-09.pdf>. [eco-art-100]
- Jonard, N. (1999). Externalités et coordination : la concurrence des standards technologiques. *Annales d'économie et de statistique*, N° 53, p149-169. Consulté le 20 octobre 2014 : <http://www.jstor.org/discover/10.2307/20076169?uid=3738016&uid=2129&uid=2&uid=70&uid=4&sid=56177598143>. [eco-art-532]
- Jondeau, E. (2001). La théorie des anticipations de la structure par terme permet-elle de rendre compte de l'évolution des taux d'intérêt sur euro-devise ? *Annales d'économie et de statistique*, 62, 139-174. Consulté le 20 octobre 2014 : <http://Annales.ensae.fr/anciens/n62/vol62-08.pdf>. [eco-art-95]
- Joutard, X., & Ruggiero, M. (2000). Recherche d'emploi et risques de récurrence du chômage: une analyse des qualifications. *Annales d'économie et de statistique*, 57, 239-265. Consulté le 20 octobre 2014 : <http://Annales.ensae.fr/anciens/n57/vol57-13.pdf>. [eco-art-292]
- Laffargue, J-P. (2000). Effets et financement d'une réduction des charges sur les bas salaires. *Revue économique*, Vol. 51, 3, 489-498. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/reco_0035-2764_2000_num_51_3_410530. [eco-art-658]
- Mahieu, R., Colin, C., & Legros, F. (2000). Des spécificités de carrière aux spécificités institutionnelles des régimes de retraite. Une simulation sur le secteur privé et la fonction publique. *Revue économique*, Vol. 51, N° hors série, 97-114. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/reco_0035-2764_2000_hos_51_1_410483. [eco-art-9]
- Malin, E., & Martimort, D. (2001). Les limites à la discrimination par les prix. *Annales d'économie et de statistique*, 62, 209-249. Consulté le 20 octobre 2014 : <http://Annales.ensae.fr/anciens/n62/vol62-11.pdf>. [eco-art-456]
- Maurin, E. (2000). Les contrats à durée déterminée et les coûts de licenciement nuisent-ils à l'embauche stable ? *Annales d'économie et de statistique*, 57, 267-291. Consulté le 20 octobre 2014 : <http://Annales.ensae.fr/anciens/n57/vol57-14.pdf>. [eco-art-221]
- Maurin, L. (2000). Les déterminants démographiques de la dette extérieure : le cas de trois pays européens, des États-Unis et du Japon. *Économie appliquée*, tome LIII, 2, 117-131. [eco-art-512]
- Mougeot, M. (2000). La tarification hospitalière : de l'enveloppe globale à la concurrence par comparaison. *Annales d'économie et de statistique*, 58, 195-213. Consulté le 20 octobre 2014 : <http://Annales.ensae.fr/anciens/n58/vol58-09.pdf>. [eco-art-411]
- Nourry, C., & Venditti, A. (2000). Agents hétérogènes, croissance et détermination de l'équilibre. *Annales d'économie et de statistique*, 59, 227-247. Consulté le 20 octobre 2014 : <http://Annales.ensae.fr/anciens/n59/vol59-11.pdf>. [eco-art-207]
- Pelgrin, F., Nauze-Fichet, E., & Dantec, A. (2000). Projection de trajectoires économiques par microsimulation. Quelle équité pour les retraites ? *Revue économique*, Vol. 51, N° hors série, 115-129. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/reco_0035-2764_2000_hos_51_1_410484. [eco-art-608]

- Rouillon, S. (2000). Catastrophe climatique irréversible et politique de l'effet de serre. *Annales d'économie et de statistique*, 59, 165-175. Consulté le 20 octobre 2014 : <http://annales.ensae.fr/anciens/n59/vol59-08.pdf>. [eco-art-648]
- Spector, D. (2001). Fédéralisme et conflit distributif. *Annales d'économie et de statistique*, 61, 151-181. Consulté le 20 octobre 2014 : <http://annales.ensae.fr/anciens/n61/vol61-10.pdf>. [eco-art-293]
- Suchère, T. (2000). Esquisse d'une herméneutique du langage comptable. *Économie appliquée*, tome LIII, 2, 133-148. [eco-art-107]
- Talon, G., Martinez, C., & Lagarenne, C. (2000). Années validées pour la retraite. Comparaison intergénérationnelle des profils d'accumulation. *Revue économique*, Vol. 51, N° hors série, 57-76. Consulté le 20 octobre 2014 : http://www.persee.fr/web/revues/home/prescript/article/reco_0035-2764_2000_hos_51_1_410481. [eco-art-141]
- Vanelle, V. (2001). L'impact de la volatilité des taux de change sur le commerce international, l'apport des études empiriques. *Économie Appliquée*, tome LIV, 2, 59-90. [eco-art-233]
- Villieu, P. (2000). Élargissement de l'Union monétaire et coordination des politiques budgétaires : un point de vue. *Annales d'économie et de statistique*, 59, 137-163. Consulté le 20 octobre 2014 : <http://annales.ensae.fr/anciens/n59/vol59-07.pdf>. [eco-art-551]
- Zenou, Y. (2000). Externalités spatiales, économies d'agglomération et formation endogène d'une ville monocentrique. *Annales d'économie et de statistique*, 58, 233-251. Consulté le 20 octobre 2014 : <http://annales.ensae.fr/anciens/n58/vol58-11.pdf>. [eco-art-272]

Annexe 1.5 : Sciences politiques

- A. Stimson, J., Tiberj, V., & Thiébaud, C. (2010). Au service de l'analyse dynamique des opinions. *Revue française de science politique*, Vol. 60, 901-926. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2010-5-page-901.htm. [spo-art-468]
- Aldrin, P. (2010). L'invention de l'opinion publique européenne. *Politix*, 89, 79-101. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2010-1-page-79.htm. [spo-art-545]
- Aust, J., & Crespy, C. (2009). Napoléon renversé ?, *Revue française de science politique*, Vol. 59, 915-938. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2009-5-page-915.htm. [spo-art-633]
- Barboni, T., & Treille, E. (2010). L'engagement 2.0. *Revue française de science politique*, Vol. 60, 1137-1157. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2010-6-page-1137.htm. [spo-art-340]
- Beauvallet, W. & Sébastien, M. (2010). L'institutionnalisation inachevée du Parlement européen. *Politix*, 89, 147-172. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2010-1-page-147.htm. [spo-art-438]
- Braconnier, C., Dormagen, J-Y. (2010). Le vote des cités est-il structuré par un clivage ethnique ?. *Revue française de science politique*, Vol. 60, 663-689. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2010-4-page-663.htm. [spo-art-160]
- Bruno, I. (2008). Y a-t-il un pilote dans l'Union ?. *Politix*, 82, 95-117. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2008-2-page-95.htm. [spo-art-335]
- Burgos, E., Mazzoleni, O., & Rayner, H. (2009). Le gouvernement de tous faute de mieux. Institutionnalisation et transformation de la « formule magique » en Suisse (1959-2003), *Politix*, 88, 39-61. Consulté le 20 octobre 2014 : <http://www.cairn.info/revue-politix-2009-4-page-39.htm>. [spo-art-410]
- Cartier, M., Coutant, I., Masclet, O., & Siblot, Y. (2010). Promotion et marginalisation des candidats de la « diversité » dans une commune de la banlieue parisienne. *Politix*, 91, 179-205. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2010-3-page-179.htm. [spo-art-146]
- Chauvin, S. (2009). En attendant les papiers. *Politix*, 87, 47-69. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2009-3-page-47.htm. [spo-art-577]

- Daniel, F.-J. (2010). La « gestionnarisation » des politiques de protection de la nature aux Pays-Bas. *Politix* 91, 157-177. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2010-3-page-157.htm. [spo-art-457]
- Delalande Émile-Justin, N. (2008). Menier, un chocolatier en République. *Politix*, 84, p9-33. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2008-4-page-9.htm. [spo-art-249]
- Denord, F., & Schwartz, A. (2010). L'économie (très) politique du traité de Rome. *Politix*, 89, 35-56. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2010-1-page-35.htm. [spo-art-637]
- Favre, P. (2009). Quand la police fabrique l'ordre social. *Revue française de science politique*, Vol. 59, 1231-1248. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2009-6-page-1231.htm. [spo-art-358]
- Frisque, C. (2010). Des espaces médiatiques et politiques locaux ?. *Revue française de science politique*, Vol. 60, 951-973. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2010-5-page-951.htm. [spo-art-389]
- Goirand, C. (2010). Penser les mouvements sociaux d'Amérique latine. *Revue française de science politique*, Vol. 60, 445-466. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2010-3-page-445.htm. [spo-art-643]
- Halpern, C., & Le Galès, P. (2011). Pas d'action publique autonome sans instruments propres. *Revue française de science politique*, Vol. 61, 51-78. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2011-1-page-51.htm. [spo-art-170]
- Hamman, P. (2008). Patrons et milieux d'affaires français dans l'arène politique et électorale (XIXe - XXe siècles) : quelle historiographie ?. *Politix*, 84, 35-59. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2008-4-page-35.htm. [spo-art-433]
- Harguindéguy, J.-B., & Cole, A. (2009). La politique linguistique de la France à l'épreuve des revendications ethnoterritoriales. *Revue française de science politique*, Vol. 59, 939-966. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2009-5-page-939.htm. [spo-art-391]
- Lefranc, S. (2007). Convertir le grand nombre à la paix... *Politix*, 80, 7-29. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2007-4-page-7.htm. [spo-art-586]
- Moreau de Bellaing, C. (2009). Violences illégitimes et publicité de l'action policière. *Politix*, 87, 119-141. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2009-3-page-119.htm. [spo-art-380]
- Petitfils, A.-S. (2010). Mobilisations et luttes internes autour des questions homosexuelles à l'UMP : l'« affaire Vanneste ». *Politix* 92, 99-124. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2010-4-page-99.htm. [spo-art-73]
- Phélippeau, E. (2010). Genèse d'une codification. *Revue française de science politique*, Vol. 60, 519-563. Consulté le 20 octobre 2014 : <http://www.cairn.info/revue-francaise-de-science-politique-2010-3-page-519.htm>. [spo-art-323]
- Ravinet, P. (2011). La coordination européenne «à la bolognaise ». *Revue française de science politique*, Vol. 61, 23-49. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2011-1-page-23.htm. [spo-art-117]
- Rea, A., Jacobs, D., Teney, C., & Delwit, P. (2010). Les comportements électoraux des minorités ethniques à Bruxelles. *Revue française de science politique*, Vol. 60, 691-717. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2010-4-page-691.htm. [spo-art-153]
- Roché, S. (2009). Politique et administration dans la formulation d'une politique publique. *Revue française de science politique*, Vol. 59, 1147-1174. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2009-6-page-1147.htm. [spo-art-241]
- Roger, A. (2010). Constructions savantes et légitimation des politiques européennes. *Revue française de science politique*, Vol. 60, 1091-1113. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2010-6-page-1091.htm. [spo-art-51]
- Roger, A. (2010). Représenter les pauvres. Construction et gestion des clientèles politiques dans une métropole roumaine. *Politix*, 92, 31-51. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2010-4-page-31.htm. [spo-art-548]

Rumpala, Y. (2009). La « consommation durable » comme nouvelle phase d'une gouvernementalisation de la consommation. *Revue française de science politique*, Vol. 59, 967-996. Consulté le 20 octobre 2014 : www.cairn.info/revue-francaise-de-science-politique-2009-5-page-967.htm. [spo-art-642]

Vigour, C. (2009). Politiques et gouvernements fédéraux en Belgique, entre contraintes coalitionnelles et logique de compromis. *Politix*, 88, 63-86. Consulté le 20 octobre 2014 : www.cairn.info/revue-politix-2009-4-page-63.htm. [spo-art-142]

Annexe 1.6 : Anthropologie

Affergan, F. (2007). L'anthropologie cognitive existe-t-elle?. *L'Homme*, 184, 85-105. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2007-4-page-85.htm. [ant-art-41]

Albert, J-P. (2009). Les animaux, les hommes et l'Alliance. *L'Homme*, 189, 81-114. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2009-1-page-81.htm. [ant-art-368]

Assayag, J. (2008). L'anthropologie en guerre. *L'Homme*, 187-188, 135-167. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2008-3-page-135.htm. [ant-art-556]

Barbier-Le Déroff, M-A. (2007). Des grands dans l'espace des petits. *Ethnologie française*, Vol. 37, 655-662. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2007-4-page-655.htm. [ant-art-183]

Chamak, B. (2009). L'autisme dans un service de pédopsychiatrie. *Ethnologie française*, Vol. 39, 425-433. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2009-3-page-425.htm. [ant-art-150]

Conord, S. (2007). Usages et fonctions de la photographie. *Ethnologie française*, Vol. 37, 11-22. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2007-1-page-11.htm. [ant-art-196]

De L'Estoile, B. (2008). L'anthropologie après les musées ?. *Ethnologie française*, Vol. 38, 665-670. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2008-4-page-665.htm. [ant-art-149]

Delaisi de Parseval, G., & Collard, C. (2007). La gestation pour autrui. *L'Homme*, 183, 29-53. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2007-3-page-29.htm. [ant-art-352]

Dupuy, F. (2009). Les monnaies primitives. *L'Homme*, 190, 129-151. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2009-2-page-129.htm. [ant-art-57]

Eideliman, J-S. (2009). Familles à l'épreuve. *Ethnologie française*, Vol. 39, 435-442. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2009-3-page-435.htm. [ant-art-218]

Gaudy, C. (2008). " Être une femme " sur un plateau de tournage. *Ethnologie française*, Vol. 38, 107-117. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2008-1-page-107.htm. [ant-art-192]

Gérard, B. (2009). De l'ethnographie à l'ethnomusicologie. *L'Homme*, 191, 139-173. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2009-3-page-139.htm. [ant-art-93]

Glassner, J-J. (2009). Essai pour une définition des écritures. *L'Homme*, 192, 7-22. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2009-4-page-7.htm. [ant-art-85]

Isnart, C. (2009). Le chant des origines. *Ethnologie française*, Vol. 39, 483-493. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2009-3-page-483.htm. [ant-art-222]

Lange, M-F. (2007). Espaces scolaires en Afrique francophone. *Ethnologie française*, Vol. 37, 639-645. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2007-4-page-639.htm. [ant-art-430]

Laplantine, F. (2007). Penser en images. *Ethnologie française*, Vol. 37, 47-56. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2007-1-page-47.htm. [ant-art-265]

Lassègue, J., Rosenthal, V., & Visetti, Y-M. (2009). Économie symbolique et phylogenèse du langage. *L'Homme*, 192, 67-100. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2009-4-page-67.htm. [ant-art-121]

Laurière, C. (2008). L'anthropologie et le politique, les prémisses. *L'Homme*, 187-188, 69-92. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2008-3-page-69.htm. [ant-art-211]

- Le Bras, H. (2007). Quelles statistiques ethniques?. *L'Homme*, 184, 7-24. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2007-4-page-7.htm. [ant-art-312]
- Macdonald (2008). L'anthropologie sociale en France, dans quel état ?. *Ethnologie française*, Vol. 38, 617-625. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2008-4-page-617.htm. [ant-art-133]
- Marchive, A. (2007). Le rituel, la règle et les savoirs. *Ethnologie française*, Vol. 37, 597-604. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2007-4-page-597.htm. [ant-art-535]
- Maresca, S. (2007). Photographes et ethnologues. *Ethnologie française*, vol. 37, p61-67. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2007-1-page-61.htm. [ant-art-592]
- Pasquier, D. (2008). Conflits professionnels et luttes pour la visibilité à la télévision française. *Ethnologie française*, Vol. 38, 23-30. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2008-1-page-23.htm. [ant-art-231]
- Perrenoud, M. (2008). Les musicos au miroir des artisans du bâtiment. *Ethnologie française*, Vol. 38, 101-106. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2008-1-page-101.htm. [ant-art-544]
- Puig, N. (2009). Amour, honte et prestige au Caire. *L'Homme*, 190, 51-77. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2009-2-page-51.htm. [ant-art-245]
- Rhani, Z. (2009). Le chérif et la possédée. *L'Homme*, 190, 27-50. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2009-2-page-27.htm. [ant-art-347]
- Sizorn, M. (2008). Une ethnologue en " Trapèze " : sport, art ou spectacle ?. *Ethnologie française*, Vol. 38, 79-88. Consulté le 20 octobre 2014 : www.cairn.info/revue-ethnologie-francaise-2008-1-page-79.htm. [ant-art-415]
- Toffin, G. (2009). Exposer/Voir. *L'Homme*, 189, 139-164. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2009-1-page-139.htm. [ant-art-452]
- Vinciguerra, L. (2009). Les trois liens anthropologiques. *L'Homme*, 191, 7-26. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2009-3-page-7.htm. [ant-art-627]
- Zonabend, F. (2007). Adopter des sœurs. *L'Homme*, 183, 9-28. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-homme-2007-3-page-9.htm. [ant-art-271]

Annexe 1.7 : Géographie

- Ait-Alhayane, K. (2010). Gestion des ressources naturelles : une démarche pour voir et comprendre l'espace des hommes. *L'Espace géographique*, Vol. 39, 20-34. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2010-1-page-20.htm. [geo-art-575]
- Boulrier, J., & Simon, L. (2010). Les forêts au secours de la planète : quel potentiel de stockage du carbone ?. *L'Espace géographique*, Tome 39, 309-324. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2010-4-page-309.htm. [geo-art-110]
- Bouna Timera, M. (2009). L'identité sénégalaise à travers la géographie scolaire, de l'indépendance aux années 1990. *L'Espace géographique*, Vol. 38, 233-250. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2009-3-page-233.htm. [geo-art-297]
- Chaléard, J.-L. , Mesclier, E. (2010). Nouveaux acteurs, nouveaux produits et expansion du marché foncier dans le nord du Pérou. *Annales de géographie*, 676, 678-696. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2010-6-page-678.htm. [geo-art-1]
- Debrie, J., & Lavaud-Letilleul, V. (2009). La décentralisation des ports en France : de la recomposition institutionnelle aux mutations fonctionnelles. L'exemple de Toulon. *Annales de géographie*, 669, 498-521. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2009-5-page-498.htm. [geo-art-560]
- Déry, S. (2008). Les aires protégées, nouvel outil du dynamisme des États sud-est asiatiques ?. *Annales de géographie*, 659, 72-92. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2008-1-page-72.htm. [geo-art-134]

- Djament-Tran, G., & Laporte, A. (2010). Comment Berlin devint capitale de l'Allemagne réunifiée. *L'Espace géographique*, Vol. 39, 146-158. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2010-2-page-146.htm. [geo-art-550]
- Ducruet, C., Gelézeau, V., & Roussin, S. (2008). Les connexions maritimes de la Corée du Nord. Recompositions territoriales dans la péninsule Coréenne et dynamiques régionales en Asie du Nord-Est. *L'Espace géographique*, Tome 37, 208-224. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2008-3-page-208.htm. [geo-art-255]
- Fleury, A. (2009). Espaces publics et environnement dans les politiques urbaines à Paris et à Berlin. *Annales de géographie*, 669, 522-542. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2009-5-page-522.htm. [geo-art-250]
- Fol, S., & Cunningham-Sabot, E. (2010). " Déclin urbain " et Shrinking Cities : une évaluation critique des approches de la décroissance urbaine. *Annales de géographie*, 674, 359-383. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2010-4-page-359.htm. [geo-art-348]
- Forriez, M., Martin, P., & Nottale, L. (2010). Lois d'échelle et transitions fractal-non fractal en géographie. *L'Espace géographique*, vol. 39, 97-112. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2010-2-page-97.htm. [geo-art-372]
- Fourault-Cauët, V. (2010). Le paysage, outil de territorialisation et d'aménagement incomplet pour les forêts méditerranéennes ?. *Annales de géographie*, 673, 268-292. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2010-3-page-268.htm. [geo-art-393]
- Germes, M., Schirmel, H., Brailich, A., Glasze, G., & Pütz, R. (2010). Les grands ensembles de banlieue comme menaces urbaines ?. *Annales de géographie*, 675, 515-535. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2010-5-page-515.htm. [geo-art-282]
- Godet, L. (2010). La " nature ordinaire " dans le monde occidental. *L'Espace géographique*, Tome 39, 295-308. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2010-4-page-295.htm. [geo-art-]
- Grasland, C., & Van Hamme, G. (2010). La relocalisation des activités industrielles : une approche centre-périphérie des dynamiques mondiale et européenne. *L'Espace géographique*, Vol. 39, 1-19. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2010-1-page-1.htm. [geo-art-475]
- Kalaora, B. (2010). Les zones humides et le Conservatoire du littoral : perceptions et cadre d'expérience. *L'Espace géographique*, Tome 39, 361-374. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2010-4-page-361.htm. [geo-art-298]
- Kouzmine, Y., Fontaine, J., Yousfi, B-E., & Otmane, T. (2009). Étapes de la structuration d'un désert : l'espace saharien algérien entre convoitises économiques, projets politiques et aménagement du territoire. *Annales de géographie*, 670, 659-685. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2009-6-page-659.htm. [geo-art-381]
- Le Blanc, A. (2010). La conservation des ruines traumatiques, un marqueur ambigu de l'histoire urbaine. *L'Espace géographique*, Tome 39, 253-266. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2010-3-page-253.htm. [geo-art-349]
- Marius-Gnanou, K. (2010). Nouvelles activités économiques et dynamique métropolitaine : le cas de la périphérie Sud de Chennai. *Annales de géographie*, 671-672, 28-51. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2010-1-page-28.htm. [geo-art-79]
- Poinsot, Y. (2009). Protection de la grande faune et territoires : deux modèles de gestion dans la cordillère Cantabrique. *L'Espace géographique*, Vol. 38, 289-302. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2009-4-page-289.htm. [geo-art-580]
- Puel, G. (2009). Géographie des lieux d'accès à Internet. Les conditions de l'accès public et les modèles d'usages dans les grandes villes de Chine. *L'Espace géographique*, Vol. 38, 17-29. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2009-1-page-17.htm. [geo-art-34]
- Quertamp, F. (2010). La périurbanisation de Hanoi. Dynamiques de la transition urbaine vietnamienne et métropolisation. *Annales de géographie*, 671-672, 93-119. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2010-1-page-93.htm. [geo-art-7]
- Raimond, C., Breton, C., Abouya, A., & Moussa, A. (2010). Planification territoriale et accès aux ressources naturelles. Retour sur la démarche participative des opérations de sécurisation foncière dans le Nord Cameroun.

- Annales de géographie*, 676, 639-656. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2010-6-page-639.htm. [geo-art-237]
- Renard, F., & Chapon, P.-M. (2010). Une méthode d'évaluation de la vulnérabilité urbaine appliquée à l'agglomération lyonnaise. *L'Espace géographique*, Vol. 39, 35-50. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2010-1-page-35.htm. [geo-art-111]
- Rozenblat, C., Bohan, C., & Benet, G. (2008). Les réseaux d'entreprises multinationales et l'attractivité des villes d'Europe Centrale. *Annales de géographie*, 664, 70-84. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2008-6-page-70.htm. [geo-art-270]
- Sohn, C. (2009). Des villes entre coopération et concurrence. Analyse des relations culturelles transfrontalières dans le cadre de "Luxembourg et Grande Région, Capitale européenne de la Culture 2007". *Annales de géographie*, 667, 228-246. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2009-3-page-228.htm. [geo-art-447]
- Taillard, C. (2009). Un exemple réussi de régionalisation transnationale en Asie orientale : les corridors de la Région du Grand Mékong. *L'Espace géographique*, Vol. 38, 1-16. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2009-1-page-1.htm. [geo-art-48]
- Terrier, E. (2009). Les mobilités spatiales des étudiants internationaux. Déterminants sociaux et articulation des échelles de mobilité. *Annales de géographie*, 670, 609-636. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2009-6-page-609.htm. [geo-art-6]
- Tsayem Demaze, M. (2010). Éviter ou réduire la déforestation pour atténuer le changement climatique : le pari de la REDD. *Annales de géographie*, 674, 338-358. Consulté le 20 octobre 2014 : www.cairn.info/revue-annales-de-geographie-2010-4-page-338.htm. [geo-art-200]
- Yvroux, C. (2009). L'impact du contexte géopolitique sur "l'habiter" des populations d'Hébron-Al Khalil (Cisjordanie). *L'Espace géographique*, Vol. 38, 222-232. Consulté le 20 octobre 2014 : www.cairn.info/revue-espace-geographique-2009-3-page-222.htm. [geo-art-655]

Annexe 1.8 : Histoire

- Allain, T. (2008). Entre fête civique et reconstruction identitaire en Hollande : Enkhuizen en 1772. *Revue d'histoire moderne et contemporaine*, 55-1, 74-97. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2008-1-page-74.htm. [his-art-363]
- Bantigny, L. (2007). Temps, âge et génération à l'épreuve de la guerre : La mémoire, l'histoire, l'oubli des appelés en Algérie. *Revue historique*, 641, 165-179. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2007-1-page-165.htm. [his-art-599]
- Baroin, C., & Valette-Cagnac, E. (2007). S'habiller et se déshabiller en Grèce et à Rome (III). *Revue historique*, 643, 517-551. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2007-3-page-517.htm. [his-art-127]
- Bonin, H. (2010). L'action du Premier ministre Chaban-Delmas pour rendre la France industrielle plus performante (1969-1972). *Revue historique*, 654, 397-426. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2010-2-page-397.htm. [his-art-460]
- Briand, J. (2010). Foi, politique et information en Champagne au XVe siècle. *Revue historique*, 653, 59-97. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2010-1-page-59.htm. [his-art-91]
- Buylaert, F. (2010). Les anciens Pays-Bas : nouvelles approches. *Revue historique*, 653, 3-25. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2010-1-page-3.htm. [his-art-184]
- Chapoutot, J. (2008). Comment meurt un Empire : le nazisme, l'Antiquité et le mythe. *Revue historique*, 647, 657-676. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2008-3-page-657.htm. [his-art-495]
- Delleaux, F. (2010). Diffusion et application des méthodes culturelles flamandes dans les anciens Pays-Bas méridionaux au xviii^e siècle. *Revue historique*, 653, 27-58. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2010-1-page-27.htm. [his-art-568]

- Depret, I. (2009). La construction nationale en Épire, 1912-1939 : le rôle du facteur confessionnel et du haut clergé orthodoxe. *Revue d'histoire moderne et contemporaine*, 56-3, 123-149. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2009-3-page-123.htm. [his-art-97]
- Durand, G., & Klein, J-F. (2010). Une impossible liaison ?. *Revue d'histoire moderne et contemporaine*, 57-1, 139-167. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2010-1-page-139.htm. [his-art-341]
- Fournier, M. (2007). La " révolution " de la lecture romanesque au XVIIIe siècle en France : institutionnalisation de la lecture et émergence d'une nouvelle sensibilité. *Revue d'histoire moderne et contemporaine*, 54-2, 55-73. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2007-2-page-55.htm. [his-art-645]
- Guzzi-Heeb, S., & Payot, C. (2010). Conflits politiques et réseaux sociaux au XVIIIe siècle. *Revue d'histoire moderne et contemporaine*, 57-4, 72-96. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2010-4-page-72.htm. [his-art-351]
- Huret, R. (2009). " Une armée de délinquants fiscaux ". *Revue d'histoire moderne et contemporaine*, 56-2, 188-210. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2009-2-page-188.htm. [his-art-17]
- Jami, C. (2008). Pékin au début de la dynastie Qing : capitale des savoirs impériaux et relais de l'Académie royale des sciences de Paris. *Revue d'histoire moderne et contemporaine*, 55-2, 43-69. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2008-2-page-43.htm. [his-art-371]
- Julerot, V. (2006). La première entrée de l'évêque : réflexions sur son origine. *Revue historique*, 639, 635-675. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2006-3-page-635.htm. [his-art-511]
- Kozovoï, A. (2009). La rencontre Brejnev-Nixon de 1972 et la culture de guerre froide soviétique. *Revue historique*, 652, 897-914. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2009-4-page-897.htm. [his-art-513]
- Lanoë, C. (2008). Images, masques et visages. *Revue d'histoire moderne et contemporaine*, 55-1, 7-27. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2008-1-page-7.htm. [his-art-598]
- Legay, M-L. (2010). Capitalisme, crises de trésorerie et donneurs d'avis : une relecture des années 1783-1789. *Revue historique*, 655, 577-608. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2010-3-page-577.htm. [his-art-166]
- Lemercier, C. (2007). Devenir une institution locale : la Chambre de commerce de Paris au XIXe siècle. *Revue d'histoire moderne et contemporaine*, 54-3, 40-62. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2007-3-page-40.htm. [his-art-489]
- Lignereux, Y. (2010). Le visage du roi, de François Ier à Louis XIV. *Revue d'histoire moderne et contemporain*, 57-4, 30-50. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2010-4-page-30.htm. [his-art-310]
- Lyon-Caen, N. (2008). La minute du miracle. *Revue historique*, 645, 61-83. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2008-1-page-61.htm. [his-art-333]
- Markovits, R. (2009). L'incendie de la comédie de Genève (1768) : Rousseau, Voltaire et l'impérialisme culturel français. *Revue historique*, 652, 831-873. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2009-4-page-831.htm. [his-art-609]
- Michaud, F. (2009). Le pauvre transformé : les hommes, les femmes et la charité à Marseille, du XIIIe siècle jusqu'à la Peste noire. *Revue historique*, 650, 243-290. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2009-2-page-243.htm. [his-art-63]
- Pattieu, S. (2010). Plein soleil, ou l'histoire du bronzage. *Revue d'histoire moderne et contemporaine*, 57-3, 160-166. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2010-3-page-160.htm. [his-art-635]
- Payen-Variéras, E. (2008). " Entre priorités marchandes et rationalité industrielle : les cadres salariés du Central Pacific Railroad, 1869-1889. *Revue d'histoire moderne et contemporaine*, 55-4, 122-159. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2008-4-page-122.htm. [his-art-205]

- Peretz, P. (2007). Un tournant humanitaire de la politique étrangère américaine ?. *Revue d'histoire moderne et contemporaine*, 54-3, 138-158. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2007-3-page-138.htm. [his-art-263]
- Regourd F. (2008). Capitale savante, capitale coloniale : sciences et savoirs coloniaux à Paris aux XVIIe et XVIIIe siècles. *Revue d'histoire moderne et contemporaine*, 55-2, 121-151. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2008-2-page-121.htm. [his-art-216]
- Tatarenko, L. (2008). Violence et luttes religieuses dans la Confédération polono-lithuanienne (fin XVIe - milieu du XVIIe siècle) : l'exemple de la confrontation entre uniates et orthodoxes. *Revue historique*, 648, 857-890. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2008-4-page-857.htm. [his-art-135]
- Vidal, G. (2006). Violence et politique dans la France des années 1930 : le cas de l'autodéfense communiste. *Revue historique*, 640, 901-922. Consulté le 20 octobre 2014 : www.cairn.info/revue-historique-2006-4-page-901.htm. [his-art-397]
- Wheeler, D. (2010). Le climat de l'océan Atlantique aux XVIIe-XVIIIe siècles selon les journaux de bord de la Marine britannique. *Revue d'histoire moderne et contemporaine*, 57-3, 42-69. Consulté le 20 octobre 2014 : www.cairn.info/revue-d-histoire-moderne-et-contemporaine-2010-3-page-42.htm. [his-art-152]

Annexe 1.9 : Sciences de l'Information et de la Communication

- Baligand, M-P., Battisti, M., Chapoy, E., Merrien, D., Haettiger, M., Cottart, O-J., & Bourrion, D. (2010). Des pratiques professionnelles renouvelées. *Documentaliste-Sciences de l'Information*, Vol. 47, 44-55. Consulté le 20 octobre 2014 : www.cairn.info/revue-documentaliste-sciences-de-l-information-2010-2-page-44.htm. [infcom-art-232]
- Bessi res, D. (2009). La d finition de la communication publique : des enjeux disciplinaires aux changements de paradigmes organisationnels. *Communication & Organisation*, 35, 14-28. Consult  le 20 octobre 2014 : www.cairn.info/revue-communication-et-organisation-2009-1-page-14.htm. [infcom-art-601]
- Beste r, E. (2010). Les services pour les archives ouvertes : de la r f rence   l'expertise. *Documentaliste-Sciences de l'Information*, Vol. 47, 4-15. Consult  le 20 octobre 2014 : www.cairn.info/revue-documentaliste-sciences-de-l-information-2010-4-page-4.htm. [infcom-art-496]
- Boub e, N. (2010). La m thode de l'autoconfrontation : une m thode bien adapt e   l'investigation de l'activit  de recherche d'information ?. * tudes de communication*, 35, 47-60. Consult  le 20 octobre 2014 : www.cairn.info/revue-etudes-de-communication-2010-2-page-47.htm. [infcom-art-473]
- Boukacem-Zeghmouri, C. (2010). Pratiques de consultation des revues  lectroniques par les enseignants chercheurs : les STM en France. *Documentaliste-Sciences de l'Information*, Vol. 47, 4-13. Consult  le 20 octobre 2014 : www.cairn.info/revue-documentaliste-sciences-de-l-information-2010-2-page-4.htm. [infcom-art-463]
- Burlat, C. (2009). La l gitimit  des transferts de pratiques de management au sein de la firme " locale-globale ". *Communication & Organisation*, 36, 70-81. Consult  le 20 octobre 2014 : www.cairn.info/revue-communication-et-organisation-2009-2-page-70.htm. [infcom-art-261]
- Carmes, M. (2009). Les activit s de contribution dans la politique intranet d'un Conseil G n ral : de l'injonction   une s mio-politique de la " transversalit  " en organisation. * tudes de communication*, 33, 19-40. Consult  le 20 octobre 2014 : www.cairn.info/revue-etudes-de-communication-2009-2-page-19.htm. [infcom-art-590]
- Combes, C. (2011). La consommation de s ries   l' preuve d'internet. *R seaux*, 165, 137-163. Consult  le 20 octobre 2014 : www.cairn.info/revue-reseaux-2011-1-page-137.htm. [infcom-art-21]
- Coutant, A. (2009). Les cordonniers sont toujours les plus mal chauss s : le mythe de Babel v cu par les professionnels de la communication. *Communication & Organisation*, 35, 226-239. Consult  le 20 octobre 2014 : www.cairn.info/revue-communication-et-organisation-2009-1-page-226.htm. [infcom-art-581]
- Delcambre, P. (2009). L'activit  d' valuation et les syst mes d'information. L' valuation est aussi un travail : langagier, assist , organis . * tudes de communication*, 33, 79-100. Consult  le 20 octobre 2014 : www.cairn.info/revue-etudes-de-communication-2009-2-page-79.htm. [infcom-art-126]

- Ferreira, J. (2010). La médiatisation : de la production à la circulation des nouvelles. *Les Enjeux de l'information et de la communication*, Vol. 2010, 121-143. Consulté le 20 octobre 2014 : www.cairn.info/revue-les-enjeux-de-l-information-et-de-la-communication-2010-1-page-121.htm. [infcom-art-366]
- Gadras, S. (2010). La médiation politique comme cadre d'analyse de l'évolution des pratiques de communication au sein de l'espace public local. *Les Enjeux de l'information et de la communication*, Dossier 2010, 12-25. Consulté le 20 octobre 2014 : www.cairn.info/revue-les-enjeux-de-l-information-et-de-la-communication-2010--page-12.htm. [infcom-art-106]
- Gkouskou-Giannakou, P., Huet, F., Choplin, H., & Lenay, C. (2009). La campagne de communication comme réponse d'un réseau à un événement ? *Communication & Organisation*, 35, 82-94. Consulté le 20 octobre 2014 : www.cairn.info/revue-communication-et-organisation-2009-1-page-82.htm. [infcom-art-203]
- Hardy, M. (2009). Quelles approches possibles du terrain chinois en communication organisationnelle? Situation et retour d'expérience. *Études de communication*, 32, 109-128. Consulté le 20 octobre 2014 : www.cairn.info/revue-etudes-de-communication-2009-1-page-109.htm. [infcom-art-604]
- Kogan, A-F. (2008). L'ancrage social de la téléassistance pour personnes âgées : des actes de communication à l'information organisationnelle. *Les Enjeux de l'information et de la communication*, Vol. 2008, 31-43. Consulté le 20 octobre 2014 : www.cairn.info/revue-les-enjeux-de-l-information-et-de-la-communication-2008-1-page-31.htm. [infcom-art-466]
- Le Cam, F. (2010). Histoires et filiations du terme 'weblog' (1992-2003). Perspectives pour penser l'histoire de certaines pratiques sociales sur le web. *Les Enjeux de l'information et de la communication*, Vol. 2010, 97-120. Consulté le 20 octobre 2014 : www.cairn.info/revue-les-enjeux-de-l-information-et-de-la-communication-2010-1-page-97.htm. [infcom-art-620]
- Le Deuff, O. (2010). La culture de l'information et l'héritage documentaire. *Documentaliste-Sciences de l'Information*, Vol. 47, 4-11. Consulté le 20 octobre 2014 : www.cairn.info/revue-documentaliste-sciences-de-l-information-2010-3-page-4.htm. [infcom-art-179]
- Le Saulnier, G. (2011). Les policiers réels devant leurs homologues fictifs : fiction impossible ? *Réseaux*, 165, 109-135. Consulté le 20 octobre 2014 : www.cairn.info/revue-reseaux-2011-1-page-109.htm. [infcom-art-5]
- Lefebvre, M. (2010). Rendre public le processus d'évaluation de la recherche. *Réseaux*, 164, 71-96. Consulté le 20 octobre 2014 : www.cairn.info/revue-reseaux-2010-6-page-71.htm. [infcom-art-432]
- Lelong, B., & Gayoso, E. (2010). Innovation avec l'usager et plateformes collaboratives. *Réseaux*, 164, 97-126. Consulté le 20 octobre 2014 : www.cairn.info/revue-reseaux-2010-6-page-97.htm. [infcom-art-455]
- Lépine, V. (2009). La reconnaissance au travail par la construction d'une relation agissante. *Communication & Organisation*, 36, 96-107. Consulté le 20 octobre 2014 : www.cairn.info/revue-communication-et-organisation-2009-2-page-96.htm. [infcom-art-547]
- Lethiais, V., & Roudaut, K. (2010). Les amitiés virtuelles dans la vie réelle. *Réseaux*, 164, 13-49. Consulté le 20 octobre 2014 : www.cairn.info/revue-reseaux-2010-6-page-13.htm. [infcom-art-398]
- Liquète, V., Fabre, I., & Gardiès, C. (2010). Faut-il reconsidérer la médiation documentaire ? *Les Enjeux de l'information et de la communication*, Dossier 2010, 43-57. Consulté le 20 octobre 2014 : www.cairn.info/revue-les-enjeux-de-l-information-et-de-la-communication-2010--page-43.htm. [infcom-art-62]
- Maurel, D. (2010). Sense-making : un modèle de construction de la réalité et d'appréhension de l'information par les individus et les groupes. *Études de communication*, 35, 31-46. Consulté le 20 octobre 2014 : www.cairn.info/revue-etudes-de-communication-2010-2-page-31.htm. [infcom-art-374]
- Merzeau, L., Le Crosnier, H., Denis, V., Picard, J-M., Juanals, B., & Mercier, S. (2010). Une nouvelle dimension de l'information. *Documentaliste-Sciences de l'Information*, Vol. 47, 32-41. Consulté le 20 octobre 2014 : www.cairn.info/revue-documentaliste-sciences-de-l-information-2010-1-page-32.htm. [infcom-art-139]
- Ogez, E., Castagnac, G., Motta, V., & Bourcet, E. (2010). Le social sous toutes ses facettes. *Documentaliste-Sciences de l'Information*, Vol. 47, 38-52. Consulté le 20 octobre 2014 : www.cairn.info/revue-documentaliste-sciences-de-l-information-2010-3-page-38.htm. [infcom-art-502]
- Perreur, N. (2011). La néo-série, arène d'évaluation culturelle d'une société américaine en crise. *Réseaux*, 165, 83-108. Consulté le 20 octobre 2014 : www.cairn.info/revue-reseaux-2011-1-page-83.htm. [infcom-art-131]
- Peyrelong, M-F. (2009). L'activité du bibliothécaire entre ombre et lumière. *Communication & Organisation*, 36, 82-94. Consulté le 20 octobre 2014 : www.cairn.info/revue-communication-et-organisation-2009-2-page-82.htm. [infcom-art-344]

Vacher, B. (2009). Articulation entre communication, information et organisation en SIC. *Les Enjeux de l'information et de la communication*, Vol. 2009, 119-143. Consulté le 20 octobre 2014 : www.cairn.info/revue-les-enjeux-de-l-information-et-de-la-communication-2009-1-page-119.htm. [infcom-art-311]

Vacher, B. (2010). Sens et normes font-ils bon ménage dans les organisations ?. *Études de communication*, 34, 127-142. Consulté le 20 octobre 2014 : www.cairn.info/revue-etudes-de-communication-2010-1-page-127.htm. [infcom-art-523]

Annexe 1.10 : Sociologie

Augé, A. (2007). Les solidarités des élites politiques au Gabon : entre logique ethno-communautaire et réseaux sociaux. *Cahiers internationaux de sociologie*, 123, 245-268. Consulté le 20 octobre 2014 : www.cairn.info/revue-cahiers-internationaux-de-sociologie-2007-2-page-245.htm. [soc-art-83]

Authier, J-V. (2008). Les citadins et leur quartier. *L'Année sociologique*, Vol. 58, 21-46. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-annee-sociologique-2008-1-page-21.htm. [soc-art-157]

Banos, V., Candau, J., & Baud, A-C. (2009). Anonymat en localité. *Cahiers internationaux de sociologie*, 127, 247-267. Consulté le 20 octobre 2014 : www.cairn.info/revue-cahiers-internationaux-de-sociologie-2009-2-page-247.htm. [soc-art-503]

Berger, N. (2010). Sociologie analytique, mécanismes et causalité : histoire d'une relation complexe. *L'Année sociologique*, Vol. 60, 419-443. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-annee-sociologique-2010-2-page-419.htm. [soc-art-253]

Berrebi-Hoffmann, I., & Grémion, P. (2009). Élitisme intellectuel et réforme de l'État. *Cahiers internationaux de sociologie*, 126, 39-59. Consulté le 20 octobre 2014 : www.cairn.info/revue-cahiers-internationaux-de-sociologie-2009-1-page-39.htm. [soc-art-56]

Bronner, G. (2006). L'acteur social est-il (déjà) soluble dans les neurosciences ?. *L'Année sociologique*, Vol. 56, 331-351. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-annee-sociologique-2006-2-page-331.htm. [soc-art-60]

Caradec, V., Lefrançois, C., & Poli, A. (2009). Quand la discrimination et la diversité se déclinent selon l'âge : émergence, appropriations, ambivalences. *Cahiers internationaux de sociologie*, 127, 223-245. Consulté le 20 octobre 2014 : www.cairn.info/revue-cahiers-internationaux-de-sociologie-2009-2-page-223.htm. [soc-art-588]

Caveng, R. (2009). Inversement des positions et ré-enchantement de l'interaction. *Actes de la recherche en sciences sociales*, 178, 88-97. Consulté le 20 octobre 2014 : www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2009-3-page-88.htm. [soc-art-64]

Commaille, J., & Dumoulin, L. (2009). Heurs et malheurs de la légalité dans les sociétés contemporaines. une sociologie politique de la "judiciarisation". *L'Année sociologique*, Vol. 59, 63-107. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-annee-sociologique-2009-1-page-63.htm. [soc-art-35]

Coulangeon, P. (2010). Les métamorphoses de la légitimité. *Actes de la recherche en sciences sociales*, 181-182, 88-105. Consulté le 20 octobre 2014 : www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2010-1-page-88.htm. [soc-art-628]

Didry, C. (2009). L'émergence du dialogue social en Europe : retour sur une innovation institutionnelle méconnue. *L'Année sociologique*, Vol. 59, 417-447. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-annee-sociologique-2009-2-page-417.htm. [soc-art-405]

Drulhe, M., Clément, S., Mantovani, J., & Membrado, M. (2007). L'expérience du voisinage : propriétés générales et spécificités au cours de la vieillesse. *Cahiers internationaux de sociologie*, 123, 325-339. Consulté le 20 octobre 2014 : www.cairn.info/revue-cahiers-internationaux-de-sociologie-2007-2-page-325.htm. [soc-art-59]

Dubey, G. (2008). Nouvelles techniques d'identification, nouveaux pouvoirs. *Cahiers internationaux de sociologie*, 125, 263-279. Consulté le 20 octobre 2014 : www.cairn.info/revue-cahiers-internationaux-de-sociologie-2008-2-page-263.htm. [soc-art-162]

- Fack, G., & Grenet, J. (2009). Sectorisation des collèges et prix des logements à Paris. *Actes de la recherche en sciences sociales*, 180, 44-62. Consulté le 20 octobre 2014 : www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2009-5-page-44.htm. [soc-art-248]
- Fillieule, R. (2008). La sociologie économique des prix contemporaine : quel apport théorique ?. *L'Année sociologique*, Vol. 58, 383-407. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-annee-sociologique-2008-2-page-383.htm. [soc-art-552]
- Geay, B. (2010). Les néo-enseignants face à l'utilitarisme. *Actes de la recherche en sciences sociales*, 184, 72-89. Consulté le 20 octobre 2014 : www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2010-4-page-72.htm. [soc-art-507]
- Grossetti, M. (2006). Trois échelles d'action et d'analyse. *L'Année sociologique*, Vol. 56, 285-307. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-annee-sociologique-2006-2-page-285.htm. [soc-art-472]
- Hamel, J. (2007). Woody allen et l'objet de la sociologie. *Cahiers internationaux de sociologie*, 123, 341-350. Consulté le 20 octobre 2014 : www.cairn.info/revue-cahiers-internationaux-de-sociologie-2007-2-page-341.htm. [soc-art-525]
- Lizé, W. (2010). Le goût jazzistique en son champ. *Actes de la recherche en sciences sociales*, 181-182, 60-87. Consulté le 20 octobre 2014 : www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2010-1-page-60.htm. [soc-art-474]
- Mennesson, C., & Clément, J-P. (2009). Boxer comme un homme, être une femme. *Actes de la recherche en sciences sociales*, 179, 76-91. Consulté le 20 octobre 2014 : www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2009-4-page-76.htm. [soc-art-557]
- Oberti, M., Sanselme, F., & Voisin, A. (2009). Ce que Sciences Po fait aux lycéens et à leurs parents : entre méritocratie et perception d'inégalités. *Actes de la recherche en sciences sociales*, 180, 102-124. Consulté le 20 octobre 2014 : www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2009-5-page-102.htm. [soc-art-177]
- Pinto, V. (2010). L'emploi étudiant et les inégalités sociales dans l'enseignement supérieur. *Actes de la recherche en sciences sociales*, 183, 58-71. Consulté le 20 octobre 2014 : www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2010-3-page-58.htm. [soc-art-118]
- Ribard, D. (2010). Professeurs, maîtres et enseignants à l'époque moderne. *Actes de la recherche en sciences sociales*, 184, 90-107. Consulté le 20 octobre 2014 : www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2010-4-page-90.htm. [soc-art-481]
- Tarrius, A. (2008). Migrations en réseaux et cohabitations urbaines aux bordures de l'Europe. *L'Année sociologique*, Vol. 58, 71-93. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-annee-sociologique-2008-1-page-71.htm. [soc-art-458]
- Vilkas, C. (2009). Des pairs aux experts : l'émergence d'un " nouveau management " de la recherche scientifique ?. *Cahiers internationaux de sociologie*, 126, 61-79. Consulté le 20 octobre 2014 : www.cairn.info/revue-cahiers-internationaux-de-sociologie-2009-1-page-61.htm. [soc-art-]. [soc-art-377]
- Villegas, M-G. (2009). Champ juridique et sciences sociales en France et aux Etats-Unis. *L'Année sociologique*, Vol. 59, 29-62. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-annee-sociologique-2009-1-page-29.htm. [soc-art-228]
- Watier, P. (2007). Les ressources de l'interprétation sociologique. *L'Année sociologique*, Vol. 57, 83-102. Consulté le 20 octobre 2014 : www.cairn.info/revue-l-annee-sociologique-2007-1-page-83.htm. [soc-art-259]
- Wieviorka, M. (2008). L'intégration : un concept en difficulté. *Cahiers internationaux de sociologie*, 125, 221-240. Consulté le 20 octobre 2014 : www.cairn.info/revue-cahiers-internationaux-de-sociologie-2008-2-page-221.htm. [soc-art-423]
- Zhang, L. (2007). Changement social et mouvements sociaux. *Cahiers internationaux de sociologie*, 122, 7-30. Consulté le 20 octobre 2014 : www.cairn.info/revue-cahiers-internationaux-de-sociologie-2007-1-page-7.htm. [soc-art-566]
- Zunigo, X. (2010). Le deuil des grands métiers. *Actes de la recherche en sciences sociales*, 184, 58-71. Consulté le 20 octobre 2014 : www.cairn.info/revue-actes-de-la-recherche-en-sciences-sociales-2010-4-page-58.htm. [soc-art-307]

Annexe 3 : Corpus de mémoires d'apprenants vietnamophones

Disciplines	Sujets de recherche	Année de soutenance	Code analyse corpus
Psychologie	Prise en charge d'un jeune enfant présentant des symptômes hyperactifs (Cas d'un garçon de 4 ans et 6 mois)	2009	pys_01
Psychologie	Etude de la relation entre mère et fils à la période d'adolescence	2009	psy_02
Economie	X	2011	eco_01
Economie	Stratégie océan bleu à la société Amazing Tour	2013	eco_02
Economie	X	2011	eco_03
Economie	Crise de dette publique de la zone euro 2010 et expériences pour le Vietnam	2011	eco_04
Economie	X	2011	eco_05
Géographie	Analyse des avantages et des inconvénients pour développer le tourisme récepteur à Ho Chi Minh ville	2008	geo_01
Géographie	Le modèle de développement du tourisme durable pour les pauvres au village de Chau Ma à la zone tampon du parc national de Cat Tien	2010	geo_02
Géographie	Une compréhension du besoin touristique des personnes âgées à Ho Chi Minh ville	2009	geo_03
Géographie	X	2009	geo_04

Géographie	Promotion de la gastronomie Vietnamiennne aux touristes	2009	geo_05
Géographie	L'état de lieu et la solution pour le développement touristique du zoo et du jardin botanique de Sai Gon	2010	geo_06
Géographie	Evaluation des effets des activités d'éducation environnementale de l'union des femmes de la ville de Danang	2008	geo_07
Géographie	Etablir le village culturel touristique BHO HỒNG de la minorité ethnique de Cờ Tu	2009	geo_08
Géographie	Vers un développement durable de l'écotourisme dans l'île de Binh Hoa Phuoc, province de Vinh Long	2007	geo_09
Linguistique	Etude des stratégies d'écoute des étudiants du département de langue et de civilisations françaises	2010	ling_01
Linguistique	Les errances dans le roman <i>étoile errante</i> de j.m.g. le clézio	2010	ling_02
Linguistique	Vers une approche pragmatique de l'emploi des pronoms d'adresse en français	2008	ling_03
Linguistique	Etudes des emprunts dans le domaine informatique	2010	ling_04
Linguistique	Le sublime et le grotesque dans "Notre Dame de Paris" de Victor Hugo	2012	ling_05
Linguistique	X	2012	ling_06

Table des matières

LISTE DES TABLEAUX	IV
LISTE DES FIGURES	VI
INTRODUCTION	1
PARTIE 1 : OBJETS ET METHODOLOGIE DE RECHERCHE	7
CHAPITRE 1. ECRITS SCIENTIFIQUES ET PHRASEOLOGIE TRANSDISCIPLINAIRE SCIENTIFIQUE	8
1.1. Ecrit scientifique vs écrit universitaire : définition.....	8
1.1.1. Écrits scientifiques	8
1.1.1.1. Définition et propriétés	8
1.1.1.2. Discours scientifique.....	9
1.1.2. Écrits universitaires	11
1.2. Phraséologie transdisciplinaire scientifique : définition	12
1.3. Phraséologie transdisciplinaire scientifique : typologie.....	15
1.3.1. Modèle de Pecman	16
1.3.2. Modèle de Riabtseva	20
1.3.3. Modèle de Granger et Paquot.....	21
1.4. Intérêt pour le lexique transdisciplinaire scientifique	23
Conclusion.....	26
CHAPITRE 2. MARQUEURS DISCURSIFS	28
2.1. Intérêt pour les marqueurs discursifs	29
2.1.1. Définition	29
2.1.2. Propriétés syntaxiques et sémantiques	31
2.2. Marqueurs discursifs : enseignement/apprentissage en FLE	33
2.2.1. Utilisation des marqueurs discursifs par des apprenants non natifs	33
2.2.2.1. Erreurs d'ordre syntaxique.....	33
2.2.2.2. Erreurs d'ordre sémantique.....	34
2.2.2.3. Erreurs d'ordre discursif	35

2.2.2.	Enseignement/apprentissage des marqueurs discursifs en FLE	37
2.2.2.4.	Les MD dans quelques manuels de FLE.....	38
2.2.2.5.	Les MD dans les grammaires de FLE.....	42
2.2.3.	Les marqueurs discursifs comme objets d'enseignement et d'apprentissage.....	44
2.3.	Marqueurs discursifs : typologie.....	46
2.3.1.	Typologie de Riegel	46
2.3.2.	Typologie de Charolles	47
2.3.3.	Typologie de Siepmann.....	49
2.3.4.	Typologie de Paillard et Vu	50
Conclusion.....		51
CHAPITRE 3. METHODOLOGIE.....		53
3.1. Une approche inscrite dans la linguistique de corpus.....		53
3.2. Constitution du corpus.....		54
3.1.1.	Choix du genre	56
3.1.2.	Sources des articles	57
3.1.3.	Périodes de rédaction des articles.....	57
3.1.4.	Langue.....	58
3.1.5.	Formats et annotations du corpus.....	58
3.1.6.	Droit d'exploitation et de reproduction des textes.....	61
3.3. Repérage des séquences polylexicales		61
3.2.1.	Extraction	63
3.2.1.1.	Des segments répétés	63
3.2.1.2.	Extraction des locutions dans les corpus analysés par Syntex	64
3.2.2.	Critères de sélection	70
3.2.2.1.	Fréquence.....	70
3.2.2.2.	Répartition disciplinaire.....	71
3.2.2.3.	Discursivité	73
3.4. Approches adoptées.....		74
Conclusion.....		76
PARTIE 2 : ANALYSES LINGUISTIQUES		77
CHAPITRE 4. TYPOLOGIE DES SEQUENCES LEXICALISEES A FONCTION DISCURSIVE		78
4.1. Principes sous-jacents de notre typologie		79
4.2. Typologie des MD.....		81
4.2.1.	Marqueurs d'enchaînement thématique	81
4.2.2.	Marqueurs de reformulation.....	87
4.2.3.	Marqueurs d'exemplification	87

4.2.4.	Marqueurs de comparaison	89
4.2.5.	Marqueurs de causalité.....	90
4.2.6.	Marqueurs de concession	91
4.2.7.	Marqueurs d'opposition	91
4.2.8.	Marqueurs de condition.....	92
4.2.9.	Marqueurs de finalité	92
4.3.	Marqueurs de reformulation	93
4.3.1.	Intérêt pour les marqueurs de reformulation	94
4.3.2.	Différents classements dans d'autres recherches.....	99
4.3.2.1.	Gülich & Kotschi (1983).....	100
4.3.2.2.	Roulet (1987)	101
4.3.2.3.	Rossari (1990, 1994).....	103
4.3.3.	Notre classement	105
Conclusion.....		107
CHAPITRE 5. MARQUEURS DE REFORMULATION : ANALYSE SYNTACTICO-SEMANTIQUE		108
5.1.	Principes de la grille d'analyse	108
5.1.1.	Objectifs	108
5.1.2.	Modèle de Paillard & Vu	109
5.2.	Paramètres de la grille d'analyse	110
5.2.1.	Définition	111
5.2.2.	Portée syntaxique	112
5.2.3.	Position syntaxique	115
5.2.4.	Valeurs sémantiques.....	117
5.3.	Clarifier	120
5.3.1.	C'est-à-dire.....	120
a.	Définition.....	122
b.	Portée syntaxique.....	122
c.	Position syntaxique.....	123
d.	Valeurs sémantiques	124
5.3.2.	À savoir	131
a.	Définition.....	131
b.	Portée syntaxique.....	132
c.	Position syntaxique.....	133
d.	Valeurs sémantiques	133
5.3.3.	Autrement dit	139
a.	Définition.....	139
b.	Portée syntaxique.....	139
c.	Position syntaxique.....	140
d.	Valeurs sémantiques	141
5.3.4.	En d'autres termes	147
a.	Définition.....	147
b.	Portée syntaxique.....	148
c.	Position syntaxique.....	148

<i>d.</i>	Valeurs sémantiques	149
5.4.	Récapituler	155
5.4.1.	En somme.....	155
<i>a.</i>	Définition.....	155
<i>b.</i>	Portée syntaxique.....	156
<i>c.</i>	Position syntaxique.....	156
<i>d.</i>	Valeurs sémantiques	157
5.4.2.	En résumé.....	162
<i>a.</i>	Définition.....	162
<i>b.</i>	Portée syntaxique.....	162
<i>c.</i>	Position syntaxique.....	163
<i>d.</i>	Valeurs sémantiques	163
5.4.3.	En un mot	168
<i>a.</i>	Définition.....	168
<i>b.</i>	Portée syntaxique.....	168
<i>c.</i>	Position syntaxique.....	168
<i>d.</i>	Valeurs sémantiques	169
5.5.	Reconsidérer.....	172
5.5.1.	Au total.....	173
<i>a.</i>	Définition.....	173
<i>b.</i>	Portée syntaxique.....	174
<i>c.</i>	Position syntaxique.....	174
<i>d.</i>	Valeurs sémantiques	175
5.5.2.	Somme toute.....	179
<i>a.</i>	Définition.....	179
<i>b.</i>	Portée syntaxique.....	179
<i>c.</i>	Position syntaxique.....	180
<i>d.</i>	Valeurs sémantiques	180
5.5.3.	En définitive	184
<i>a.</i>	Définition.....	184
<i>b.</i>	Portée syntaxique.....	184
<i>c.</i>	Position syntaxique.....	185
<i>d.</i>	Valeurs sémantiques	186
5.5.4.	En fin de compte	190
<i>a.</i>	Définition.....	190
<i>b.</i>	Portée syntaxique.....	190
<i>c.</i>	Position syntaxique.....	191
<i>d.</i>	Valeurs sémantiques	191
5.5.5.	Au fond	194
<i>a.</i>	Définition.....	194
<i>b.</i>	Portée syntaxique.....	195
<i>c.</i>	Position syntaxique.....	195
<i>d.</i>	Valeurs sémantiques	195
Conclusion.....		200
PARTIE 3 : RÉFLEXIONS DIDACTIQUES		203

CHAPITRE 6. PROPOSITIONS DIDACTIQUES POUR L'ENSEIGNEMENT/ APPRENTISSAGE DE SEQUENCES LEXICALISEES A FONCTION DISCURSIVE.....	204
6.1. Corpus de mémoires d'apprenants	205
6.1.1. Élaboration	205
6.1.2. Analyse.....	206
6.2. Linguistique de corpus et enseignement/apprentissage du français	212
6.2.1. Apprentissage sur corpus	213
6.2.2. Stratégies d'enseignement et rôle des enseignants	216
6.3. Réflexions didactiques : vers des propositions	219
6.3.1. Réflexions didactiques	220
6.3.1.1. <i>Approche inductive et approche onomasiologique</i>	220
6.3.1.2. <i>Approche déductive et approche sémasiologique</i>	234
6.3.2. Propositions didactiques.....	241
6.3.2.1. <i>Exercices d'entraînement</i>	243
6.3.2.2. <i>Fiches pédagogiques</i>	249
Fiche pédagogique 1	249
Fiche pédagogique 3	266
Fiche pédagogique 4	271
Conclusion.....	276
CONCLUSION GENERALE.....	278
BIBLIOGRAPHIE	283
INDEX.....	293
ANNEXE	295
Annexe 1: Typologie des séquences lexicalisées à fonction discursive	296
Annexe 2 : Liste des revues en SHS	305
Annexe 3 : Corpus de mémoires d'apprenants vietnamophones	324